

## DISK ARRAY CONTROLLER, ITS DISK ARRAY CONTROL UNIT AND ITS EXPANDING METHOD

**Publication number:** JP2001256003

**Publication date:** 2001-09-21

**Inventor:** FUJIMOTO KAZUHISA; KANAI HIROKI; FUJIBAYASHI AKIRA; SAKURAI WATARU

**Applicant:** HITACHI LTD

**Classification:**

- International: G06F12/08; G06F3/06; G06F12/16; G06F12/08;  
G06F3/06; G06F12/16; (IPC1-7): G06F3/06; G06F12/08;  
G06F12/16

- European: G06F3/06D

**Application number:** JP20000072469 20000310

**Priority number(s):** JP20000072469 20000310

**Also published as:**

EP1132805 (A2)

US6477619 (B1)

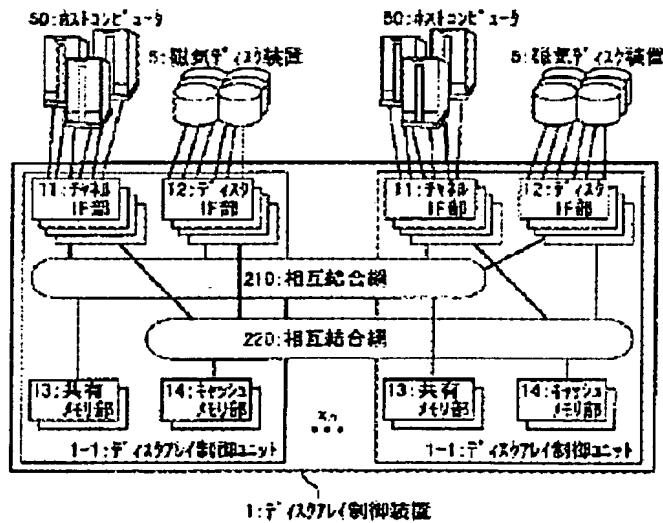
EP1132805 (A3)

[Report a data error here](#)

### Abstract of JP2001256003

**PROBLEM TO BE SOLVED:** To solve a problem that lowering of performance in data transition among plural disk array controllers is inevitable when the plural disk array controllers are attempted to be operated as one disk array controller. **SOLUTION:** This disk array controller provided with a channel IF part, a disk IF part, a cache memory part and a shared memory part and plural disk array control units to read/write data, has a mutual coupling network to connect the shared memories in the plural disk array control units and a mutual coupling network to connect the cache memory parts in the plural disk array control units are provided.

図 1



Data supplied from the esp@cenet database - Worldwide



(19)日本国特許庁 (JP)

## (12) 公開特許公報 (A)

(11)特許出願公開番号

特開2001-256003

(P2001-256003A)

(43)公開日 平成13年9月21日(2001.9.21)

(51) Int.Cl. <sup>7</sup> G 06 F 3/06 12/08	識別記号 3 0 2 5 4 0 3 2 0	F I G 06 F 3/06 12/08	テ-マゴ-ド*(参考) 3 0 2 Z 5 B 0 0 5 5 4 0 5 B 0 1 8 J 5 B 0 6 5 G 3 2 0
---	---------------------------------	-----------------------------	--

審査請求 未請求 請求項の数11 OL (全 33 頁) 最終頁に続く

(21)出願番号 特願2000-72469(P2000-72469)	(71)出願人 000005108 株式会社日立製作所 東京都千代田区神田駿河台四丁目6番地
(22)出願日 平成12年3月10日(2000.3.10)	(72)発明者 藤本 和久 東京都国分寺市東恋ヶ窪一丁目280番地 株式会社日立製作所中央研究所内
	(72)発明者 金井 宏樹 東京都国分寺市東恋ヶ窪一丁目280番地 株式会社日立製作所中央研究所内
	(74)代理人 100068504 弁理士 小川 勝男 (外1名)
	最終頁に続く

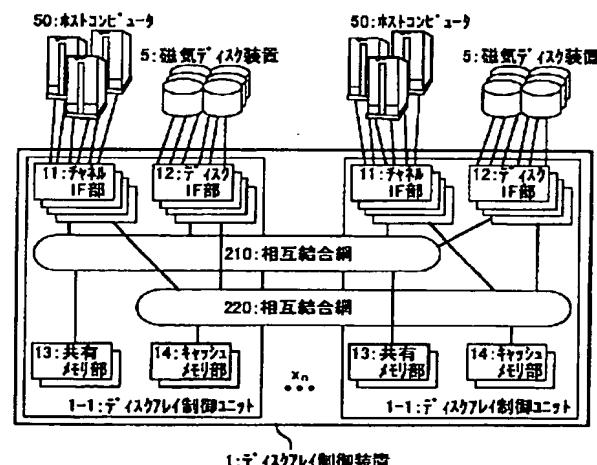
## (54)【発明の名称】 ディスクアレイ制御装置、そのディスクアレイ制御ユニットおよびその増設方法

## (57)【要約】

【課題】複数台のディスクアレイ制御装置を1つのディスクアレイ制御装置として運用しようとする場合、複数のディスクアレイ制御装置間でのデータ移行における性能低下が避けられなかった。

【課題を解決するための手段】チャネルIF部と、ディスクIF部と、キャッシュメモリ部と、共有メモリ部とを有し、データのリード/ライトを行うディスクアレイ制御ユニットを、複数ユニット有するディスクアレイ制御装置において、複数のディスクアレイ制御ユニット内の共有メモリ部間を接続する相互結合網と、複数のディスクアレイ制御ユニット内のキャッシュメモリ部間を接続する相互結合網を有する。

図 1



(2) 001-256003 (P2001-0%03)

**【特許請求の範囲】**

【請求項1】ホストコンピュータとのインターフェースを有する1または複数のチャネルインターフェース部と、ディスク装置とのインターフェースを有する1または複数のディスクインターフェース部と、前記ディスク装置に対しリード／ライトされるデータを一時的に格納するキャッシュメモリ部と、前記チャネルインターフェース部及び前記ディスクインターフェース部と前記キャッシュメモリ部との間のデータ転送に関する制御情報及び前記ディスク装置の管理情報を格納する共有メモリ部とを有し、前記チャネルインターフェース部は前記ホストコンピュータからのデータのリード／ライト要求に対し、前記ホストコンピュータとのインターフェースと前記キャッシュメモリ部との間のデータ転送を実行し、前記ディスクインターフェース部は、前記ディスク装置と前記キャッシュメモリ部との間のデータ転送を実行することにより、データのリード／ライトを行うディスクアレイ制御ユニットを、複数ユニット有するディスクアレイ制御装置であって、前記各ディスクアレイ制御ユニット内の前記チャネルインターフェース部および前記ディスクインターフェース部と前記キャッシュメモリ部との接続部間に第1の相互結合網を介して複数の前記ディスクアレイ制御ユニットにまたがって接続され、前記各ディスクアレイ制御ユニット内の前記チャネルインターフェースおよび前記ディスクインターフェースと前記共有メモリ部との接続部間に第2の相互結合網を介して複数の前記ディスクアレイ制御ユニットにまたがって接続されていることを特徴とするディスクアレイ制御装置。

【請求項2】前記各ディスクアレイ制御ユニット内の前記チャネルインターフェース部及び前記ディスクインターフェース部と前記キャッシュメモリ部は互いに前記第1の相互結合網を介して接続され、前記チャネルインターフェース部及び前記ディスクインターフェース部と前記共有メモリ部は互いに前記第2の相互結合網を介して接続されることを特徴とする請求項1記載のディスクアレイ制御装置。

【請求項3】前記各ディスクアレイ制御ユニット内の前記チャネルインターフェース部及び前記ディスクインターフェース部と前記のキャッシュメモリ部との間は、前記第1の相互結合網によって接続され、前記チャネルインターフェース部及び前記ディスクインターフェース部と前記共有メモリ部との間は、前記各ディスクアレイ制御ユニット内では直接接続されていることを特徴とする請求項1記載のディスクアレイ制御装置。

【請求項4】前記各ディスクアレイ制御ユニット内の前記チャネルインターフェース部及び前記ディスクインターフェース部と前記キャッシュメモリ部との間は、前記各ディスクアレイ制御ユニット内では、自ユニット内の第3の相互接続網を通して接続され、前記各チャネルインターフェース部及び前記複数のディスクインターフェ

ース部と前記複数の共有メモリ部との間は、前記ディスクアレイ制御ユニット内では、直接接続されていることを特徴とする請求項1記載のディスクアレイ制御装置。

【請求項5】ホストコンピュータとのインターフェースを有する1または複数のチャネルインターフェース部と、ディスク装置とのインターフェースを有する1または複数のディスクインターフェース部と、前記ディスク装置に対しリード／ライトされるデータを一時的に格納するキャッシュメモリ部と、前記チャネルインターフェース部及び前記ディスクインターフェース部と前記キャッシュメモリ部との間のデータ転送に関する制御情報及び前記ディスク装置の管理情報を格納する共有メモリ部とを有し、前記チャネルインターフェース部は前記ホストコンピュータからのデータのリード／ライト要求に対し、前記ホストコンピュータとのインターフェースと前記キャッシュメモリ部との間のデータ転送を実行し、前記ディスクインターフェース部は、前記ディスク装置と前記キャッシュメモリ部との間のデータ転送を実行することにより、データのリード／ライトを行うディスクアレイ制御ユニットであって、前記チャネルインターフェース部および前記ディスクインターフェース部と前記キャッシュメモリ部との接続部に他の前記ディスクアレイ制御ユニットへの接続のための第1の接続バスを有し、前記チャネルインターフェース部および前記ディスクインターフェース部と前記共有メモリ部との接続部に他の前記ディスクアレイ制御ユニットへの接続のための第2の接続バスを有することを特徴とするディスクアレイ制御ユニット。

【請求項6】前記ディスクアレイ制御ユニットは1つの筐体に認められ、前記第1、第2の接続バスはコネクタを備えたことを特徴とする請求項5記載のディスクアレイ制御ユニット。

【請求項7】ホストコンピュータとのインターフェースを有する1または複数のチャネルインターフェース部と、ディスク装置とのインターフェースを有する1または複数のディスクインターフェース部と、前記ディスク装置に対しリード／ライトされるデータを一時的に格納するキャッシュメモリ部と、前記チャネルインターフェース部及び前記ディスクインターフェース部と前記キャッシュメモリ部との間のデータ転送に関する制御情報及び前記ディスク装置の管理情報を格納する共有メモリ部とを有し、前記チャネルインターフェース部は前記ホストコンピュータからのデータのリード／ライト要求に対し、前記ホストコンピュータとのインターフェースと前記キャッシュメモリ部との間のデータ転送を実行し、前記ディスクインターフェース部は、前記ディスク装置と前記キャッシュメモリ部との間のデータ転送を実行することにより、データのリード／ライトを行うディスクアレイ制御ユニットであって、前記チャネルインターフェース部および前記ディスクインターフェース部と前記キャッシュメモリ部との接続部に他の前記ディスクアレイ制御ユニットへの接続のための第1の接続バスを有し、前記チャネルインターフェース部および前記ディスクインターフェース部と前記共有メモリ部との接続部に他の前記ディスクアレイ制御ユニットへの接続のための第2の接続バスを有することを特徴とするディスクアレイ制御ユニット。

(3) 001-256003 (P2001-ch7#03)

ヤッシュメモリ部との接続部に他の前記ディスクアレイ制御ユニットへの接続のための第1の接続バスを有し、前記チャネルインターフェース部および前記ディスクインターフェース部と前記共有メモリ部との接続部に他の前記ディスクアレイ制御ユニットへの接続のための第2の接続バスを有するディスクアレイ制御ユニットが複数個、および前記ディスクアレイ制御ユニットのそれぞれの第1、第2の接続バスが接続され、前記複数個のディスクアレイ制御ユニットを互いに結合し一つのディスクアレイ制御装置とせしめるスイッチボックスとを備えたことを特徴とするディスクアレイ制御装置。

【請求項8】ホストコンピュータとのインターフェースを有する1または複数のチャネルインターフェース部と、ディスク装置とのインターフェースを有する1または複数のディスクインターフェース部と、前記ディスク装置に対しリード／ライトされるデータを一時的に格納するキャッシュメモリ部と、前記チャネルインターフェース部及び前記ディスクインターフェース部と前記キャッシュメモリ部との間のデータ転送に関する制御情報及び前記ディスク装置の管理情報を格納する共有メモリ部とを有し、前記チャネルインターフェース部は前記ホストコンピュータからのデータのリード／ライト要求に対し、前記ホストコンピュータとのインターフェースと前記キャッシュメモリ部との間のデータ転送を実行し、前記ディスクインターフェース部は、前記ディスク装置と前記キャッシュメモリ部との間のデータ転送を実行することにより、データのリード／ライトを行うディスクアレイ制御ユニットであって、前記チャネルインターフェース部および前記ディスクインターフェース部と前記キャッシュメモリ部との接続部に他の前記ディスクアレイ制御ユニットへの接続のための第1の接続バスを有し、前記チャネルインターフェース部および前記ディスクインターフェース部と前記共有メモリ部との接続部に他の前記ディスクアレイ制御ユニットへの接続のための第2の接続バスを有するディスクアレイ制御ユニット、および前記ディスクアレイ制御ユニットの第1、第2の接続バスが接続され、前記複数個のディスクアレイ制御ユニットを互いに結合し一つのディスクアレイ制御装置とせしめるスイッチボックスとを備えたディスクアレイ制御装置において、前記第1、第2の接続バスおよび前記スイッチボックスにコネクタを有し、前記スイッチボックスは複数の前記ディスクアレイ制御ユニットの第1、第2の接続バスのコネクタが接続されるよう構成され、ケーブルによって前記第1、第2の接続バスのコネクタと前記スイッチボックスのコネクタを接続することにより前記ディスクアレイ制御ユニットを増設することを特徴とするディスクアレイ制御ユニットの増設方法。

【請求項9】前記スイッチボックスは前記スイッチボックスを通るデータのあて先のアドレスと接続すべき前記ディスクアレイ制御ユニットを特定する情報との対応を

示す対応情報を格納するテーブルを持ち、前記ディスクアレイ制御ユニットを増設するのに伴い、前記テーブルに増設された前記ディスクアレイ制御ユニットに関する有効な前記対応情報を追加することを特徴とする請求項8記載のディスクアレイ制御ユニットの増設方法。

【請求項10】前記ディスクアレイ制御ユニットはそれぞれ筐体に収められ、ある筐体に前記スイッチボックスと前記ディスクアレイ制御ユニットを収納し、前記スイッチボックスに複数のケーブルを接続することにより他のディスクアレイ制御ユニットを収納した筐体を接続することを特徴とする請求項8のディスクアレイ制御ユニットの増設方法。

【請求項11】ホストコンピュータとのインターフェースを有する1または複数のチャネルインターフェース部と、ディスク装置とのインターフェースを有する1または複数のディスクインターフェース部と、前記ディスク装置に対しリード／ライトされるデータを一時的に格納するキャッシュメモリ部と、前記チャネルインターフェース部及び前記ディスクインターフェース部と前記キャッシュメモリ部との間のデータ転送に関する制御情報及び前記ディスク装置の管理情報を格納する共有メモリ部とを有し、前記チャネルインターフェース部は前記ホストコンピュータからのデータのリード／ライト要求に対し、前記ホストコンピュータとのインターフェースと前記キャッシュメモリ部との間のデータ転送を実行し、前記ディスクインターフェース部は、前記ディスク装置と前記キャッシュメモリ部との間のデータ転送を実行することにより、データのリード／ライトを行うディスクアレイ制御ユニットであって、前記チャネルインターフェース部および前記ディスクインターフェース部と前記キャッシュメモリ部との接続部に他の前記ディスクアレイ制御ユニットへの接続のための第1の接続バスを有し、前記チャネルインターフェース部および前記ディスクインターフェース部と前記共有メモリ部との接続部に他の前記ディスクアレイ制御ユニットへの接続のための第2の接続バスを有する複数のディスクアレイ制御ユニット、および前記ディスクアレイ制御ユニットの第1、第2の接続バスが接続され、前記複数個のディスクアレイ制御ユニットを互いに結合し一つのディスクアレイ制御装置とせしめるスイッチボックスとが1つの筐体に収められていることを特徴とするディスクアレイ制御装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、データを複数の磁気ディスク装置に格納するディスクアレイ装置の制御装置に関する。

【0002】

【従来の技術】半導体記憶装置を記憶媒体とするコンピュータの主記憶のI/O性能に比べて、磁気ディスクを記憶媒体とするディスクサブシステム（以下「サブシス

(4) 001-256003 (P2001-> 姉

テム」という。)のI/O性能は3~4桁程度小さく、従来からこの差を縮めること、すなわちサブシステムのI/O性能を向上させる努力がなされている。サブシステムのI/O性能を向上させるための1つの方法として、複数の磁気ディスク装置でサブシステムを構成し、データを複数の磁気ディスク装置に格納する、いわゆるディスクアレイと呼ばれるシステムが知られている。

【0003】例えば、従来技術では、図2に示すようにホストコンピュータ50とディスクアレイ制御装置2との間のデータ転送を実行する複数のチャネルIF部11と、磁気ディスク装置5とディスクアレイ制御装置2間のデータ転送を実行する複数のディスクIF部12と、磁気ディスク装置5のデータを一時的に格納するキャッシュメモリ部14と、ディスクアレイ制御装置2に関する制御情報(例えば、チャネルIF部11及びディスクIF部12とキャッシュメモリ部14との間のデータ転送制御に関する情報、磁気ディスク装置5に格納するデータの管理情報)を格納する共有メモリ部13とを備え、1つのディスクアレイ制御装置2内において、共有メモリ部13およびキャッシュメモリ部14は全てのチャネルIF部11及びディスクIF部12からアクセス可能な構成となっている。このディスクアレイ制御装置2では、チャネルIF部11及びディスクIF部12と共有メモリ部13との間、及びチャネルIF部11及びディスクIF部12とキャッシュメモリ部14との間は、相互結合網21、及び相互結合網22でそれぞれ接続される。

【0004】チャネルIF部11は、ホストコンピュータ50と接続するためのインターフェース及びホストコンピュータ50に対する入出力を制御するマイクロプロセッサ(図示せず)を有している。また、ディスクIF部12は、磁気ディスク装置5と接続するためのインターフェース及び磁気ディスク装置5に対する入出力を制御するマイクロプロセッサ(図示せず)を有している。また、ディスクIF部12は、RAID機能の実行も行う。

【0005】この従来のディスクアレイ制御装置2では、1つのディスクアレイ制御装置2への接続可能なディスクの容量には上限があり、1つのディスクアレイ制御装置2がサポートできる記録データ量以上のデータを記録する必要がある場合、ディスクアレイ制御装置2を複数台設置し、ホストコンピュータ50から複数のディスクアレイ制御装置2にチャネルを接続していた。

【0006】また、1つのディスクアレイ制御装置2に接続できるホストチャネル数以上のホストコンピュータ50を接続する必要がある場合、ディスクアレイ制御装置2を複数台設置し、それぞれにホストコンピュータ50を接続していた。

【0007】また、複数のディスクアレイ制御装置2間でデータを移行させる場合、1つのホストコンピュータ50からデータの移行を行う2つのディスクアレイ制御

装置2の両方にチャネルを接続し、ホストコンピュータ50を介して、データの移行を行っていた。

【0008】また、米国特許5,680,640号に開示されている従来技術では、図3に示すように2つのディスクアレイ制御装置3間でデータの移行を行う場合、2つのディスクアレイ制御装置3の間で、ホストコンピュータ50とのインターフェースの一部(図では2本)をデータ移行バス8で接続し、一方のディスクアレイ制御装置3に繋がる磁気ディスク装置5のデータを、データ移行バス8を介して、もう一方のディスクアレイ制御装置3に繋がる磁気ディスク装置5に移行していた。

【0009】また、他の従来技術では、1つのディスクアレイ制御装置がサポートする記録データ量以上のデータを記録する必要がある場合、また1つのディスクアレイ制御装置に接続できるホストチャネル数以上のホストコンピュータを接続する必要がある場合、また、複数のディスクアレイ制御装置間でデータを移行させる場合、図4に示すようにディスクアレイ制御装置4を複数台設置し、それらのホストコンピュータ50とのインターフェースをスイッチから構成される相互結合網23を介して、ホストコンピュータ50に接続していた。

【0010】特開平11-66693号公報には、2つのディレクタ装置が共有メモリを共用して、ディスクアレイを構成するデータスピンドルが制御不能になったとき復元するアレイディスク処理装置が示されている。ここではアレイディスクを複数台設置するような構成は示されていない。

【0011】

【発明が解決しようとする課題】銀行、証券、電話会社等に代表される大企業では、従来各所に分散していたコンピュータ及びストレージを、データセンターの中に集中化してコンピュータシステム及びストレージシステムを構成することにより、コンピュータシステム及びストレージシステムの運用、保守、管理に要する費用を削減する傾向にある。

【0012】このような傾向の中で、大型/ハイエンドのディスクアレイ制御装置には、数百台以上のホストコンピュータへ接続するためのチャネルインターフェースのサポート(コネクティビティ)、数百テラバイト以上の記憶容量のサポートが要求されている。

【0013】一方、近年のオープン市場の拡大、今後予想されるストレージ・エリア・ネットワーク(SAN)の普及により、大型/ハイエンドのディスクアレイ制御装置と同様の高機能・高信頼性を備えた小規模構成(小型筐体)のディスクアレイ制御装置への要求が高まっている。

【0014】前者の要求に対しては、従来の大型/ハイエンドのディスクアレイ制御装置を複数接続して超大規模なディスクアレイ制御装置を構成する方法が考えられる。

## (5) 001-256003 (P2001-ch5蝎娃

【0015】また後者の要求に対しては、従来の大型／ハイエンドのディスクアレイ制御装置の最小構成のモデル、(例えばチャネルIF部、ディスクIF部の数を少なくしたもの)において筐体を小型化した装置を構成する方法が考えられる。また、この小型化した装置を複数台接続することにより、従来のディスクアレイ制御装置がサポートしている中規模から大規模の構成をサポートする装置を構成する方法が考えられる。

【0016】ディスクアレイ制御装置では、上記のように、小規模な構成から超大規模な構成まで、同一の高機能・高信頼性アーキテクチャで対応可能な、スケーラビリティのある構成の制御装置が必要となってきており、そのためには、複数のディスクアレイ制御装置で、1つのディスクアレイ制御装置として運用できるディスクアレイ制御装置が必要となる。

【0017】しかし、図2に示す従来技術では、ディスクアレイ制御装置2の台数を単純に増やすことによってホストコンピュータ50に接続するチャネル数、記憶容量を増やすことが可能であるが、1つのホストコンピュータ50が複数のディスクアレイ制御装置2にデータを格納する必要がある場合、ホストコンピュータ50は全てのディスクアレイ制御装置2に対してチャネルを接続する必要があった。また、データにアクセスする場合、ホストコンピュータ50はアクセスするデータがどのディスクアレイ制御装置2に接続された磁気ディスク装置5に格納されているかを把握し、目的のディスクアレイ制御装置2を特定してアクセスを行う必要があった。

【0018】したがって、複数台のディスクアレイ制御装置を1つのディスクアレイ制御装置として運用することは困難であった。

【0019】図3に示す従来技術では、データ移行バス8によりディスクアレイ制御装置間3が接続されているため、ホストコンピュータ50は1台のディスクアレイ制御装置3に接続するだけで、他のディスクアレイ制御装置3に繋がる磁気ディスク装置5に格納されたデータにアクセスすることが可能であり、複数のディスクアレイ制御装置3を1つのディスクアレイ制御装置として運用することが可能であった。

【0020】しかし、ホストコンピュータ50からディスクアレイ制御装置3に、例えばデータのリード要求があり、そのディスクアレイ制御装置3に接続された磁気ディスク装置5にデータが無かった場合、データ移行バス8を介して、他のディスクアレイ制御装置3にそのリード要求を送り、該当するデータを格納している磁気ディスク装置5が接続されたディスクアレイ制御装置3から、データ移行バス8を介して要求データを受け取り、ホストコンピュータ50へ要求データを返す必要があった。したがって、あるホストコンピュータ50が、自信が繋がるディスクアレイ制御装置3以外のディスクアレイ制御装置3に繋がる磁気ディスク装置5に格納された

データにアクセスする場合、著しく性能が低下するという問題があった。

【0021】また、上記問題を防ぐため、あるホストコンピュータ50からアクセスされる頻度の高いデータの中で、該ホストコンピュータが繋がるディスクアレイ制御装置とは異なるディスクアレイ制御装置に接続された磁気ディスク装置に格納されているデータを、予め該ホストコンピュータ50が繋がるディスクアレイ制御装置3に接続された磁気ディスク装置5に移行させておく場合、データ移行バス8を介してデータを移行させるため、性能が低いという問題があった。

【0022】また、図4に示す従来技術では、スイッチを用いた相互結合網23を介して、ホストコンピュータ50は全てのディスクアレイ制御装置4にアクセスすることが可能である。

【0023】しかし、複数のディスクアレイ制御装置4を1つのディスクアレイ制御装置として運用するためには、相互結合網23を構成するスイッチ内に、そのスイッチに接続された全てのディスクアレイ制御装置4のデータが、どのディスクアレイ制御装置4に格納されているかを示すマップを持つ必要があり、ホストコンピュータ50からアクセス要求があった場合、スイッチにおいてコマンドを解析し、要求データを格納しているディスクアレイ制御装置4に割り振る機能が必要となる。

【0024】この場合、従来のチャネルIF部11でのコマンド解析に加え、その上に繋がるスイッチにおいてもコマンドを解析する必要があるため、ホストコンピュータ50がディスクアレイ制御装置4に直接接続されている場合に比べ、性能が低下するという問題がある。

【0025】ハイエンドのディスクアレイ制御装置には以下のようないくつかの機能があった。すなわち、ある業務用のデータセット(論理ボリュームに対応)の複製を保持しており、通常の業務ではオリジナルデータセットと複製データセットに対して同時にデータの更新を行い、例えば、そのデータセットのバックアップをとる要求があった場合、複製データセットについてデータの更新を中止し、それをバックアップ用に使用し、オリジナルデータセットでは業務を継続し、バックアップが終了した時点で、オリジナルデータセットと複製のデータセットの整合性をとるという機能がある。

【0026】図2～図4に示す従来技術において、上記機能を実現する場合、異なるディスクアレイ制御装置間でデータセットの複製を保持しようとした場合には、ディスクアレイ制御装置間でデータセットを移行させる必要があるため、性能が著しく低下するという問題があった。

【0027】本発明の目的は、小規模な構成から超大規模な構成まで、同一の高機能・高信頼性アーキテクチャで対応可能な、スケーラビリティのある構成のディスクアレイ制御装置を提供することにある。

:(6) 001-256003 (P2001-EE03

【0028】より具体的には、本発明の目的は、複数台のディスクアレイ制御装置を1つのディスクアレイ制御装置として運用しようとする場合、複数のディスクアレイ制御装置間でのデータ移行による性能低下を抑えることが可能なディスクアレイシステムを提供すること、また、ディスクアレイ制御装置が有する機能を、性能低下を抑えて複数台のディスクアレイ制御装置で実現することにある。

#### 【0029】

【課題を解決するための手段】上記目的は、ホストコンピュータとのインターフェースを有する1または複数のチャネルインターフェース部と、ディスク装置とのインターフェースを有する1または複数のディスクインターフェース部と、前記ディスク装置に対しリード／ライトされるデータを一時的に格納するキャッシュメモリ部と、前記チャネルインターフェース部及び前記ディスクインターフェース部と前記キャッシュメモリ部との間のデータ転送に関する制御情報及び前記ディスク装置の管理情報を格納する共有メモリ部とを有し、前記チャネルインターフェース部及び前記ディスクインターフェース部と前記キャッシュメモリ部が接続され、前記チャネルインターフェース部及び前記ディスクインターフェース部と前記共有メモリ部が接続され、前記ホストコンピュータからのデータのリード／ライト要求に対し、前記チャネルインターフェース部は、前記ホストコンピュータとのインターフェースと前記キャッシュメモリ部との間のデータ転送を実行し、前記ディスクインターフェース部は、前記ディスク装置と前記キャッシュメモリ部との間のデータ転送を実行することにより、データのリード／ライトを行うそれ自体ディスク制御装置としての機能を備えているディスクアレイ制御ユニットを、複数ユニット有するディスクアレイ制御装置であって、前記複数のディスクアレイ制御ユニット内の前記共有メモリ部間を接続する手段と、前記複数のディスクアレイ制御ユニット内の前記キャッシュメモリ部間を接続する手段を有し、前記ディスクアレイ制御ユニット内の前記チャネルインターフェース部及び前記ディスクインターフェース部から、他の前記ディスクアレイ制御ユニット内の前記共有メモリ部のデータ、または前記キャッシュメモリ部のデータをリード／ライト可能であることを特徴とするディスクアレイ制御装置により達成される。

#### 【0030】

【発明の実施の形態】以下、大容量のデータの記憶装置として磁気ディスク装置を例にとって説明するが、大容量記憶装置として磁気ディスクに限られるものではなく、例えばDVDのような大容量記憶装置であって良い。

【0031】本発明の実施の形態の1つとして、好ましくは、前記複数のディスクアレイ制御ユニット内の前記複数のチャネルインターフェース部及び前記複数のディ

スクインターフェース部と前記複数のキャッシュメモリ部との間は、前記複数のディスクアレイ制御ユニット間に跨るスイッチを用いた相互結合網によって接続し、前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記複数の共有メモリ部との間は、前記複数のディスクアレイ制御ユニット間に跨るスイッチを用いた相互結合網によって接続する。

【0032】また、好ましくは、前記複数のディスクアレイ制御ユニット内の前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記複数のキャッシュメモリ部との間は、前記複数のディスクアレイ制御ユニット間に跨るスイッチを用いた相互結合網によって接続し、前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記複数の共有メモリ部との間は、前記ディスクアレイ制御ユニット内では直接接続し、前記ディスクアレイ制御ユニット間では、前記共有メモリ部間を前記複数のディスクアレイ制御ユニット間に跨るスイッチを用いた相互結合網によって接続する。

【0033】また、好ましくは、前記複数のディスクアレイ制御ユニット内の前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記複数のキャッシュメモリ部との間は、前記ディスクアレイ制御ユニット内では、自ユニット内のスイッチを用いた相互接続網によって接続し、前記複数のディスクアレイ制御ユニット間では、前記キャッシュメモリ部間を前記複数のディスクアレイ制御ユニット間に跨るスイッチを用いた相互結合網によって接続し、前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記複数の共有メモリ部との間は、前記ディスクアレイ制御ユニット内では、直接接続し、前記複数のディスクアレイ制御ユニット間では、前記共有メモリ部間を前記複数のディスクアレイ制御ユニット間に跨るスイッチを用いた相互結合網によって接続する。

【0034】また、ホストコンピュータと磁気ディスク装置とのデータのリード／ライトに着目すれば以下のようである。ホストコンピュータとのインターフェースを有するチャネルインターフェース部と、磁気ディスク装置とのインターフェースを有するディスクインターフェース部と、前記磁気ディスク装置に対しリード／ライトされるデータを一時的に格納するキャッシュメモリ部と、前記チャネルインターフェース部及び前記ディスクインターフェース部と前記キャッシュメモリ部との間のデータ転送に関する制御情報及び前記磁気ディスク装置の管理情報を格納する共有メモリ部と、前記チャネルインターフェース部及び前記ディスクインターフェース部と前記キャッシュメモリ部を接続する手段と、前記チャネルインターフェース部及び前記ディスクインターフェース部と前記共有メモリ部を接続する手段とを有し、前記ホストコンピュータからのデータのリード／ライト要

(7) 001-256003 (P2001->娃

求に対し、前記チャネルインターフェース部は、前記ホストコンピュータとのインターフェースと前記キャッシュメモリ部との間のデータ転送を実行し、前記ディスクインターフェース部は、前記磁気ディスク装置と前記キャッシュメモリ部との間のデータ転送を実行することにより、データのリード／ライトを行うディスクアレイ制御ユニットを、複数ユニット有するディスクアレイ制御装置であって、前記複数のディスクアレイ制御ユニット内の前記共有メモリ部間を接続する手段と、前記複数のディスクアレイ制御ユニット内の前記キャッシュメモリ部間を接続する手段を有し、該接続手段を介して、1つの前記ディスクアレイ制御ユニットにのみ接続されている前記ホストコンピュータから、該ディスクアレイ制御ユニットとは異なる前記ディスクアレイ制御ユニットにのみ接続されている前記磁気ディスク装置のデータをリード／ライト可能である。

【0035】その他、本願が開示する課題、及びその解決方法は、発明の実施形態の欄及び図面により明らかにされる。

【0036】以下、本発明の実施例を図面を用いて説明する。

【0037】[実施例1] 図1、図7、及び図8に、本発明の一実施例を示す。以下の実施例において、相互結合網はスイッチを利用したものを例に説明してあるが、相互に接続され制御情報やデータが転送されればよいのであり、例えばバスで構成されても良い。

【0038】図1に示すように、ディスクアレイ制御装置1は複数のディスクアレイ制御ユニット1-1から構成される。ディスクアレイ制御ユニット1-1は、ホストコンピュータ50とのインターフェース部（チャネルIF部）11と、磁気ディスク装置5とのインターフェース部（ディスクIF部）12と、共有メモリ部13と、キャッシュメモリ部14を有し、チャネルIF部11及びディスクIF部12と共有メモリ部13の間は複数のディスクアレイ制御ユニット1-1に跨る相互結合網210を介して接続され、チャネルIF部11及びディスクIF部12とキャッシュメモリ部14の間は複数のディスクアレイ制御ユニット1-1に跨る相互結合網220を介して接続されている。すなわち、相互結合網210、あるいは相互結合網220を介して、全てのチャネルIF部11及びディスクIF部12から、全ての共有メモリ部13、あるいは全てのキャッシュメモリ部14へアクセス可能な構成となっている。

【0039】1つのディスクアレイ制御ユニットは1つの筐体として構成されるか、またはモジュールとして構成されても良いが、それ自体1つのディスク制御装置としての機能を備えているものである。図7ではそれぞれ別個な筐体を構成しているものとして説明する。

【0040】1つのディスクアレイ制御ユニット1-1内の具体的な一例を図7に示す。ディスクアレイ制御ユ

ニット1-1は、ホストコンピュータ50との2つのチャネルIF部11と、磁気ディスク装置5との2つのディスクIF部12と、2つの共有メモリバス用スイッチ（SM-SW）110と、2つのキャッシュメモリ用バススイッチ（CM-SW）111と、2つの共有メモリ部13と、2つのキャッシュメモリ部14と、共有メモリ（SM）アクセスバス135と、136と、キャッシュメモリ（CM）アクセスバス137と、138と、筐体間SMバス141と、筐体間CMバス142を有する。筐体間SMバスはディスクアレイ制御ユニット間SMバスと、筐体間CMバスはディスクアレイユニット間CMバスと言うことが出来る。

【0041】チャネルIF部11は、ホストコンピュータ50との2つのIF（ホストIF）102と、ホストコンピュータ50に対する入出力を制御する2つのマイクロプロセッサ101と、共有メモリ部13へのアクセスを制御するアクセス制御部（SMアクセス制御部）104と、キャッシュメモリ部14へのアクセスを制御するアクセス制御部（CMアクセス制御部）105を有し、ホストコンピュータ50とキャッシュメモリ部14間のデータ転送、及びマイクロプロセッサ101と共有メモリ部13間の制御情報の転送を実行する。マイクロプロセッサ101及びホストIF102は内部バス106によって接続され、CMアクセス制御部105は2つのホストIF102に直接接続されている。また、SMアクセス制御部104は2つのマイクロプロセッサ101に直接接続されている。

【0042】ディスクIF部12は、磁気ディスク装置5との2つのIF（ドライブIF）103と、磁気ディスク装置5に対する入出力を制御する2つのマイクロプロセッサ101と、共有メモリ部13への1つのアクセス制御部（SMアクセス制御部）104と、キャッシュメモリ部14への1つのアクセス制御部（CMアクセス制御部）105を有し、磁気ディスク装置5とキャッシュメモリ部14間のデータ転送、及びマイクロプロセッサ101と共有メモリ部13間の制御情報の転送を実行する。マイクロプロセッサ101及びドライブIF103は内部バス106によって接続され、CMアクセス制御部105は2つのドライブIF103に直接接続されている。また、SMアクセス制御部104は2つのマイクロプロセッサ101に直接接続されている。ディスクIF部はRAID機能の実行も行う。

【0043】共有メモリ部13は、共有メモリ（SM）コントローラ107とメモリモジュール109とを有し、ディスクアレイ制御ユニット1-1の制御情報（例えば、チャネルIF部11及びディスクIF部12とキャッシュメモリ部14との間のデータ転送制御に関する情報、磁気ディスク装置5に記録するデータの管理情報）等を格納する。

【0044】キャッシュメモリ部14は、キャッシュメ

!(8) 001-256003 (P2001-ch墨娃

モリ(CM)コントローラ108とメモリモジュール109を有し、磁気ディスク装置5へ記録するデータを一時的に格納する。

【0045】SMアクセス制御部104には2本のSMアクセスバス135を接続し、それらを2つの異なるSM-SW110にそれぞれ接続する。SM-SW110には2本のアクセスバス136を接続し、それらを2つの異なるSMコントローラ107にそれぞれ接続する。したがってSMコントローラ107には、2つのSM-SW110から1本ずつ、計2本のアクセスバス136が接続される。こうすることにより、1つのSMアクセス制御部104から1つのSMコントローラ107へのアクセスルートが2つとなる。これにより、1つのアクセスバスまたはSM-SW110に障害が発生した場合でも、もう1つのアクセスルートにより共有メモリ部13へアクセスすることが可能となるため、耐障害性を向上させることができる。

【0046】CMアクセス制御部105には2本のCMアクセスバス137を接続し、それらを2つの異なるCM-SW111にそれぞれ接続する。CM-SW111には2本のアクセスバス138を接続し、それらを2つの異なるCMコントローラ108にそれぞれ接続する。したがってCMコントローラ108には、2つのCM-SW111から1本ずつ、計2本のアクセスバス138が接続される。こうすることにより、1つのCMアクセス制御部105から1つのCMコントローラ108へのアクセスルートが2つとなる。これにより、1つのアクセスバスまたはCM-SW111に障害が発生した場合でも、もう1つのアクセスルートによりキャッシュメモリ部14へアクセスすることが可能となるため、耐障害性を向上させることができる。

【0047】SM-SW110には、2つのチャネルIF部11と、2つのディスクIF部12からそれぞれ1本ずつ、計4本のSMアクセスバス135が接続される。また、SM-SW110には、2つの共有メモリ部13へのアクセスバス136が1本ずつ、計2本接続される。また、他のディスクアレイ制御ユニット1-1のSM-SW110と接続するための筐体間SMバス141が2本接続される。これは一方が入力、他方が出力用であっても良く、また、それそれが双方向の情報の転送が出来るものであっても良い。

【0048】CM-SW111には、2つのチャネルIF部11と、2つのディスクIF部12からそれぞれ1本ずつ、計4本のCMアクセスバス137が接続される。また、CM-SW111には、2つのキャッシュメモリ部14へのアクセスバス138が1本ずつ、計2本接続される。また、他のディスクアレイ制御ユニット1-1のCM-SW111と接続するための筐体間CMバス142が2本接続される。この2本もまた筐体間SMバスと同様の性質を持っている。

【0049】SM-SW110、あるいはCM-SW111では、上記のようなアクセスバスが接続されるため、SM-SW110内、あるいはCM-SW111内では、チャネルIF部11及びディスクIF部12からの4本のアクセスバスからの要求を、自ディスクアレイ制御ユニット1-1内の共有メモリ部13、あるいはキャッシュメモリ部14への2本のアクセスバスと、他ディスクアレイ制御ユニット1-1内の共有メモリ部13、あるいはキャッシュメモリ部14への2本の筐体間アクセスバスに振り分ける機能を有する。

【0050】図7でSM-SW110はチャネルIF部11およびディスクIF部12と共有メモリ部13との接続部であり、CM-SW111はチャネルIF部11およびディスクIF部12とキャッシュメモリ部14との接続部である。

【0051】図7に示すディスクアレイ制御ユニット1-1を2台接続したディスクアレイ制御装置1の例を図8に示す。

【0052】各ディスクアレイ制御ユニット1-1-1、1-1-2内のSM-SW110間は、各筐体間SMバス141を筐体間SM-SW121を介して接続する。

【0053】同様に、各ディスクアレイ制御ユニット1-1-1、1-1-2内のCM-SW111間は、各筐体間CMバス142を筐体間CM-SW122を介して接続する。SW121、122はスイッチボックスとして実装される。

【0054】本実施例のように、接続するディスクアレイ制御ユニット1-1が2台の場合は、筐体間SW121、122を介さずに直接筐体間バスを接続しても本発明を実施する上で問題は無い。そうすることにより、筐体間SW121、122で発生するデータ転送処理のオーバヘッドを削減することが可能となり、性能が向上する。

【0055】筐体間SW121、122に接続できるバス数を、図8に示す例より増やすことにより、3台以上のディスクアレイ制御ユニット1-1を接続することが可能となる。また、筐体間SW121、122に実装可能なバス数には物理的な限界があるため、接続するディスクアレイ制御ユニット1-1を増やしていく場合、1つの筐体間SWだけでは接続できなくなる。そうした場合は、筐体間SWを多段に接続することで、接続可能なディスクアレイ制御ユニット1-1の台数を増やすことが可能となる。スイッチボックスにより3台のディスクアレイ制御ユニットを相互結合した実装例を図19に示す。詳細は後述する。

【0056】また、3台以上のディスクアレイ制御ユニット1-1を接続する場合、図13に示すように各ディスクアレイ制御ユニット1-1内のSM-SW110及びCM-SW111間をループ状に接続する。このよう

(9) 001-256003 (P2001-d版03)

な構成では、筐体間SW121、122を用いずに複数のディスクアレイ制御ユニット1-1を接続することが可能となる。

【0057】この場合には各ディスクアレイ制御ユニット1-1から出ている筐体間SMバス同士、筐体間CMバス同士をコネクタで接続すれば良い。図では省略してあるが筐体間SMバス141、および筐体間CMバス142のディスクアレイ制御ユニット1-1の外の部分にコネクタを設ければディスクアレイ制御ユニットを増設するのに都合が良い。図8でSM-SW110、筐体間SMバス141および筐体間SM-SW121で図1の相互結合網210を構成し、CM-SW111、筐体間CMバス142および筐体間CM-SW122で図1の相互結合網220を構成する。

【0058】図8において、ディスクアレイ制御装置1-1-1に接続されたホストコンピュータ50からデータを読み出す場合の一例を述べる。

【0059】まず、ホストコンピュータ50は、自身が接続されているディスクアレイ制御装置1-1-1内のチャネルIF部11にデータの読み出し要求を発行する。要求を受けたチャネルIF部11内のマイクロプロセッサ101は、2つのディスクアレイ制御ユニット1-1-1、1-1-2内の共有メモリ部13の両方にアクセスし、要求されたデータがどの磁気ディスク装置5内に格納されているかを調べる。共有メモリ部13には、要求データのアドレスとそのデータがどの磁気ディスク装置5に格納されているかを対応させる変換テーブルが格納されており、要求されたデータがどの磁気ディスク装置5内に格納されているかを調べることができる。

【0060】次に、要求を受けたチャネルIF部11内のマイクロプロセッサ101は、2つのディスクアレイ制御ユニット1-1-1、1-1-2内の共有メモリ部13の両方にアクセスし、各ディスクアレイ制御ユニット1-1-1、1-1-2内のキャッシュメモリ部14に要求されたデータが格納されているかどうかを確認する。共有メモリ部13にはキャッシュメモリ部14内データのディレクトリ情報が格納されており、各キャッシュメモリ部14に要求データが存在するかどうかを確認できる。

【0061】それによりディスクアレイ制御ユニット1-1-1のキャッシュメモリ部14内にデータがあった場合は、そのデータを自身のCM-SW111を介してチャネルIF部11まで転送し、ホストコンピュータ50に送る。ディスクアレイ制御ユニット1-1-2のキャッシュメモリ部14内にデータがあった場合は、そのデータをディスクアレイ制御ユニット1-1-2内のCM-SW111、筐体間CM-SW122、自身のCM-SW111を介してチャネルIF部11まで転送し、ホストコンピュータ50に送る。

【0062】どのキャッシュメモリ部14内にもデータ

が存在しなかった場合、マイクロプロセッサ101は要求データが格納されている磁気ディスク装置5が接続されているディスクIF部12内のマイクロプロセッサ101に対し、要求データを読み出し、キャッシュメモリ部14に格納するように命令を発行する。命令を受けたディスクIF部12内のマイクロプロセッサ101は、要求データが格納されている磁気ディスク装置5からデータを読み出し、2つのディスクアレイ制御ユニット1-1-1、1-1-2の内のどちらか一方のキャッシュメモリ部14に要求データを格納する。要求データが格納されている磁気ディスク装置5が繋がるディスクアレイ制御ユニット1-1-2内のキャッシュメモリ部14にデータを格納する場合は、そのディスクアレイ制御ユニット1-1-2内のCM-SW111を介することにより、また、要求データが格納されている磁気ディスク装置5が繋がるディスクアレイ制御ユニット1-1-2とは異なるディスクアレイ制御ユニット1-1-1内のキャッシュメモリ部14にデータを格納する場合は、それぞれのCM-SW111及び筐体間CM-SW122を介することにより、キャッシュメモリ部14へデータが転送される。

【0063】ディスクIF部12内のマイクロプロセッサ101は、要求データをキャッシュメモリ部14へ格納した後、命令を発行したチャネルIF部11内のマイクロプロセッサ101に、データを格納したキャッシュメモリ部14を伝える。それを受けたチャネルIF部11内のマイクロプロセッサ101は、データが格納されたキャッシュメモリ部14からデータを読み出し、ホストコンピュータ50へ送る。

【0064】本実施例によれば、ホストコンピュータ50は要求データがどのディスクアレイ制御ユニット1-1に繋がる磁気ディスク装置5に格納されているかを意識することなく、自身が繋がるディスクアレイ制御ユニット1-1にアクセス要求を発行するだけで、データの書き込み及び読み出しを行うことが可能になり、ホストコンピュータ50に対して、複数台のディスクアレイ制御ユニット1-1を1つのディスクアレイ制御装置1に見せることが可能となる。

【0065】また、要求を受けたディスク制御ユニット1-1とは異なるディスクアレイ制御ユニット1-1に繋がる磁気ディスク装置5からデータを読み出す場合、内部の相互結合網及びキャッシュメモリ部14を介して、データを読み出すことが出来るため、両方のディスクアレイ制御ユニット1-1のチャネルIF部11を介してデータを移行する必要が無くなり、データの読み出し及び書き込みの性能の低下を抑えることが可能となる。

【0066】[実施例2] 図5、図9、及び図10に、本発明の他の実施例を示す。

【0067】図5に示すように、複数のディスクアレイ制御ユニット1-2からなるディスクアレイ制御装置1

(専0) 01-256003 (P2001-d慮\$複姓

の構成は、チャネル I F 部 1 1 及びディスク I F 部 1 2 と共有メモリ部 1 3 の間の接続構成を除いて、実施例 1 の図 1 に示す構成と同様である。

【0068】チャネル I F 部 1 1 及びディスク I F 部 1 2 と共有メモリ部 1 3 の間は、ディスクアレイ制御ユニット 1 - 2 内において直接に接続されている。また、複数のディスクアレイ制御ユニット 1 - 2 の間では、共有メモリ部 1 3 は相互結合網 2 4 を介して接続されている。

【0069】上記のように、この実施例ではディスクアレイ制御ユニット 1 - 2 内においてチャネル I F 部 1 1 及びディスク I F 部 1 2 と共有メモリ部 1 3 を直接接続することにより、実施例 1 で示した相互結合網 2 1 0 を介して接続する場合に比べ、共有メモリ部 1 3 へのアクセス時間を短縮することが可能になる。

【0070】1 つのディスクアレイ制御ユニット 1 - 2 内の具体的な一例を図 9 に示す。

【0071】ディスクアレイ制御ユニット 1 - 2 内の構成も、チャネル I F 部 1 1 及びディスク I F 部 1 2 と共有メモリ部 1 3 の間の接続構成を除いて、実施例 1 の図 7 に示す構成と同様である。

【0072】ディスクアレイ制御ユニット 1 - 2 は、ホストコンピュータ 50 との 2 つのチャネル I F 部 1 1 と、磁気ディスク装置 5 との 2 つのディスク I F 部 1 2 と、2 つのキャッシュメモリ用バススイッチ (CM-SW) 1 1 1 と、2 つの共有メモリ部 1 3 と、2 つのキャッシュメモリ部 1 4 と、共有メモリ (SM) アクセスバス 1 3 9 と、キャッシュメモリ (CM) アクセスバス 1 3 7 と、CM アクセスバス 1 3 8 と、筐体間 SM パス 1 4 3 と、筐体間 CM パス 1 4 2 を有する。

【0073】SM アクセス制御部 1 0 4 には 2 本の SM アクセスバス 1 3 9 を接続し、それらを 2 つの異なる SM コントローラ 1 0 7 にそれぞれ接続する。したがって SM コントローラ 1 0 7 には、2 つのチャネル I F 部 1 1 及び 2 つのディスク I F 部 1 2 から 1 本ずつ、計 4 本の SM アクセスバス 1 3 9 が接続される。また、他のディスクアレイ制御ユニット 1 - 2 の SM コントローラ 1 0 7 と接続するための筐体間 SM パス 1 4 3 が 2 本接続される。

【0074】SM コントローラ 1 0 7 では、上記のようなアクセスバスが接続されるため、SM コントローラ 1 0 7 内では、チャネル I F 部 1 1 及びディスク I F 部 1 2 からの 4 本の SM アクセスバス 1 3 9 からの要求を、メモリモジュール 1 0 9 へのアクセスバスと、他ディスクアレイ制御ユニット 1 - 2 内の共有メモリ部 1 3 への 2 本の筐体間 SM アクセスバス 1 4 3 に分ける機能を有する。

【0075】図 9 に示すディスクアレイ制御ユニット 1 - 2 を 2 台接続したディスクアレイ制御装置 1 の例を図 10 に示す。

【0076】各ディスクアレイ制御ユニット 1 - 2 内の共有メモリ部 1 3 間は、各筐体間 SM パス 1 4 3 を筐体間 SM-SW 1 2 1 を介して接続する。

【0077】それ以外は、実施例 1 の図 8 に示す構成と同様である。この場合は SM コントローラ 1 0 7 がチャネル I F 部 とディスク I F 部 と共有メモリ部との接続部をなす。

【0078】実施例 1 と同様に、接続するディスクアレイ制御ユニット 1 - 2 が 2 台の場合は、筐体間 SW 1 2 1 、 1 2 2 を介さずに直接筐体間バスを接続しても本発明を実施する上で問題は無い。そうすることにより、筐体間 SW 1 2 1 、 1 2 2 で発生するデータ転送処理のオーバヘッドを削減することが可能となり、性能が向上する。

【0079】また、実施例 1 と同様に、筐体間 SW 1 2 1 、 1 2 2 に接続できるバス数を、図 10 に示す例により増やすことにより、3 台以上のディスクアレイ制御ユニット 1 - 2 を接続することが可能となる。また、筐体間 SW 1 2 1 、 1 2 2 に実装可能なバス数には物理的な限界があるため、接続するディスクアレイ制御ユニット 1 - 2 を増やしていく場合、1 つの筐体間 SW だけでは接続できなくなる。そうした場合は、筐体間 SW を多段に接続することで、接続可能なディスクアレイ制御ユニット 1 - 2 の台数を増やすことが可能となる。これは実施例 1 と同様スイッチボックスとして実装される。

【0080】また、3 台以上のディスクアレイ制御ユニット 1 - 2 を接続する場合、実施例 1 において、図 13 に示した構成と同様のループ状の構成をとることにより、筐体間 SW 1 2 1 、 1 2 2 を用いずに複数のディスクアレイ制御ユニット 1 - 2 を接続することが可能となる。

【0081】本実施例において、ホストコンピュータ 50 から磁気ディスク装置 5 へのデータの読み出し / 書き込みを行う場合の、ディスクアレイ制御ユニット 1 - 2 内での各部の動作は、他のディスクアレイ制御ユニット 1 - 2 内の共有メモリ部 1 3 へアクセスが、自ディスクアレイ制御ユニット 1 - 2 内の共有メモリ部 1 3 及び相互結合網 2 4 を介して行われることを除いて、実施例 1 と同様である。なお、筐体間 SM パス 1 4 3 、筐体間 SM-SW 1 2 1 、および SM コントローラ 1 0 7 で相互結合網 2 4 が構成される。

【0082】本実施例によれば、ホストコンピュータ 50 は要求データがどのディスクアレイ制御ユニット 1 - 2 に繋がる磁気ディスク装置 5 に格納されているかを意識することなく、自身が繋がるディスクアレイ制御ユニット 1 - 2 にアクセス要求を発行するだけで、データの書き込み及び読み出しを行うことが可能になり、ホストコンピュータ 50 に対して、複数台のディスクアレイ制御ユニット 1 - 2 を 1 つのディスクアレイ制御装置 1 に見せることが可能となる。

## (単1))01-256003 (P2001-E焼娃

【0083】また、要求を受けたディスク制御ユニット1-2とは異なるディスクアレイ制御ユニット1-2に繋がる磁気ディスク装置5からデータを読み出す場合、内部の相互結合網及びキャッシュメモリ部14を介して、データを読み出すことが出来たため、両方のディスクアレイ制御ユニット1-2のチャネルIF部11を介してデータを移行する必要が無くなり、データの読み出し及び書き込みの性能の低下を抑えることが可能となる。

【0084】[実施例3]図6、図11、及び図12に、本発明の他の実施例を示す。

【0085】図6に示すように、複数のディスクアレイ制御装置1-3からなるディスクアレイ制御装置1の構成は、チャネルIF部11及びディスクIF部12とキャッシュメモリ部14の間の接続構成を除いて、実施例2の図5に示す構成と同様である。

【0086】チャネルIF部11及びディスクIF部12とキャッシュメモリ部14の間は、ディスクアレイ制御ユニット1-3内において相互結合網22を介して接続されている。また、複数のディスクアレイ制御ユニット1-3の間では、キャッシュメモリ部14は相互結合網25を介して接続されている。実施例2と同様にチャネルIF部11およびディスクIF部12と共有メモリ部13が直接接続されているのに対し、チャネルIF部11およびディスクIF部12とキャッシュメモリ部14との間は相互結合網22を介して接続されているのは次の理由である。共有メモリ部13で扱われる制御情報は例えば数バイトであるのに対し、キャッシュメモリ部14では例えば数Kバイトの単位でデータが扱われる。そのため、相互結合網22を通して限られたビン数で接続しながらスループットを上げることが図られる。

【0087】上記のように、チャネルIF部11及びディスクIF部12とキャッシュメモリ部14を接続する相互結合網22と、ディスクアレイ制御ユニット1-3間でキャッシュメモリ部14を接続する相互結合網24を分離することにより、異なるディスクアレイ制御ユニット1-3のキャッシュメモリ部14間でデータを移行させる場合、ホストコンピュータ50からのアクセス要求時のキャッシュメモリ部14へのアクセスを妨げることがなくなる。キャッシュメモリ間のデータの移行はディスクIF部が司る。異なるディスクアレイ制御ユニット1-3のキャッシュメモリ部14間でデータを移行させる機能は、1つのディスクアレイ制御ユニット1-3に対してホストコンピュータ50からのアクセスが集中した場合、アクセス頻度の少ないディスクアレイ制御ユニットにデータを移動させるために必要な機能である。

【0088】1つのディスクアレイ制御ユニット1-3内の具体的な一例を図11に示す。

【0089】ディスクアレイ制御ユニット1-3内の構成も、チャネルIF部11及びディスクIF部12とキャッシュメモリ部14の間の接続構成を除いて、実施例

2の図9に示す構成と同様である。ここではCMコントローラ108がチャネルIF部とディスクIF部とキャッシュメモリ部との接続部をなす。

【0090】ディスクアレイ制御ユニット1-3は、ホストコンピュータ50との2つのチャネルIF部11と、磁気ディスク装置5との2つのディスクIF部12と、2つのキャッシュメモリ用バススイッチ(CM-SW)111と、2つの共有メモリ部13と、2つのキャッシュメモリ部14と、共有メモリ(SM)アクセスパス139と、キャッシュメモリ(CM)アクセスパス137と、CMアクセスパス138と、筐体間SMバス143と、筐体間CMバス144を有する。

【0091】CMアクセス制御部105には2本のCMアクセスパス137を接続し、それらを2つの異なるCM-SW111にそれぞれ接続する。CM-SW111には2本のアクセスパス138を接続し、それらを2つの異なるCMコントローラ108にそれぞれ接続する。したがってCMコントローラ108には、2つのCM-SW111から1本ずつ、計2本のアクセスパス138が接続される。また、他のディスクアレイ制御ユニット1-3のCMコントローラ108と接続するための筐体間CMバス144が2本接続される。

【0092】CMコントローラ108では、上記のようなアクセスパスが接続されるため、CMコントローラ108内では、CM-SW111からの2本のCMアクセスパス138からの要求を、メモリモジュール109へのアクセスパスと、他ディスクアレイ制御ユニット1-3内のキャッシュメモリ部14への2本の筐体間CMアクセスパス144に振分ける機能を有する。

【0093】CM-SW111には、2つのチャネルIF部11と、2つのディスクIF部12からそれぞれ1本ずつ、計4本のCMアクセスパス137が接続される。また、CM-SW111には、2つのキャッシュメモリ部14へのアクセスパス138が1本ずつ、計2本接続される。

【0094】CM-SW111では、上記のようなアクセスパスが接続されるため、CM-SW111内では、チャネルIF部11及びディスクIF部12からの4本のCMアクセスパス137からの要求を、キャッシュメモリ部14への2本のCMアクセスパス138に振分ける機能を有する。

【0095】図11に示すディスクアレイ制御ユニット1-3を2台接続したディスクアレイ制御装置1の例を図12に示す。

【0096】各ディスクアレイ制御ユニット1-3内のキャッシュメモリ部14間は、各筐体間CMバス144を筐体間CM-SW122を介して接続する。それ以外は、実施例2の図10に示す構成と同様である。

【0097】実施例2と同様に、接続するディスクアレイ制御ユニット1-3が2台の場合は、筐体間SW12

(注2) 01-256003 (P2001-違法)

1、122を介さずに直接筐体間バスを接続しても本発明を実施する上で問題は無い。そうすることにより、筐体間SW121、122で発生するデータ転送処理のオーバヘッドを削減することが可能となり、性能が向上する。

【0098】また、実施例2と同様に、筐体間SWに接続できるバス数を、図12に示す例より増やすことにより、3台以上のディスクアレイ制御ユニット1-3を接続することが可能となる。また、筐体間SW121、122に実装可能なバス数には物理的な限界があるため、接続するディスクアレイ制御ユニット1-3を増やしていく場合、1つの筐体間SWだけでは接続できなくなる。そうした場合は、筐体間SWを多段に接続することで、接続可能なディスクアレイ制御ユニット121、122の台数を増やすことが可能となる。

【0099】また、3台以上のディスクアレイ制御ユニット1-3を接続する場合、実施例1において、図13に示した構成と同様のループ状の構成をとることにより、筐体間SW121、122を用いずに複数のディスクアレイ制御ユニット1-3を接続することが可能となる。

【0100】本実施例において、ホストコンピュータ50から磁気ディスク装置5へのデータの読み出し／書き込みを行う場合の、ディスクアレイ制御ユニット1-3内の各部の動作は、他のディスクアレイ制御ユニット1-3内のキャッシュメモリ部14へアクセスが、自ディスクアレイ制御ユニット1-3内のキャッシュメモリ部14及び相互結合網25を介して行われることを除いて、実施例2と同様である。

【0101】本実施例によれば、ホストコンピュータ50は要求データがどのディスクアレイ制御ユニット1-3に繋がる磁気ディスク装置5に格納されているかを意識することなく、自身が繋がるディスクアレイ制御ユニット1-3にアクセス要求を発行するだけで、データの書き込み及び読出しを行うことが可能になり、ホストコンピュータ50に対して、複数台のディスクアレイ制御ユニット1-3を1つのディスクアレイ制御装置1に見せることが可能となる。

【0102】また、要求を受けたディスク制御ユニット1-3とは異なるディスクアレイ制御ユニット1-3に繋がる磁気ディスク装置5からデータを読み出す場合、内部の相互結合網及びキャッシュメモリ部14を介して、データを読み出すことが出来るため、両方のディスクアレイ制御ユニット1-3のチャネルIF部11を介してデータを移行する必要が無くなり、データの読出し及び書き込みの性能の低下を抑えることが可能となる。

【0103】次に本発明の実施例の利用例について説明する。ハイエンドのディスクアレイ制御装置には以下のような機能がある。すなわち、ある業務用のデータセット(論理ボリュームに対応)の複製を保持しておき、通常

の業務ではオリジナルデータセットと複製データセットに対して同時にデータの更新を行い、例えば、そのデータセットのバックアップをとる要求があった場合、複製データセットについてデータの更新を中止し、それをバックアップ用に使用し、オリジナルデータセットは業務を継続し、バックアップが終了した時点で、オリジナルデータセットと複製のデータセットのデータの整合性をとるという機能がある。

【0104】実施例1に示すディスクアレイ制御装置1において、異なるディスクアレイ制御ユニット1-1間でデータセットの複製を保持する場合に、上記機能を実現する方法について、図8を用いて説明する。

【0105】ここでは仮に、図8に示すディスクアレイ制御ユニット1-1-1に繋がる磁気ディスク装置5にオリジナルのデータセットが格納され、ディスクアレイ制御ユニット1-1-2に繋がる磁気ディスク装置5に複製のデータセットが格納されているとする。また、ディスクアレイ制御ユニット1-1-1に繋がるホストコンピュータ50が通常の業務を行い、ディスクアレイ制御ユニット1-1-2に繋がるホストコンピュータ50が自身に繋がるテープ装置(図示していない)にデータのバックアップをとる作業を行うこととする。

【0106】通常の業務では、ディスクアレイ制御ユニット1-1-1に繋がるホストコンピュータ50から目的のデータセットへデータの書き込み要求があった場合、ディスクアレイ制御ユニット1-1-1に繋がるホストコンピュータ50に繋がるチャネルIF部11内のマイクロプロセッサ101は、ホストコンピュータ50から送られてきたデータをディスクアレイ制御ユニット1-1-1のキャッシュメモリ部14に転送し、書き込む。次に、前記のマイクロプロセッサ101は、オリジナルのデータセットが格納されている磁気ディスク装置5が繋がるディスクIF部12内のマイクロプロセッサ101に、ディスクアレイ制御ユニット1-1-1の共有メモリ部13を介して命令を発行し、ディスクアレイ制御ユニット1-1-1のキャッシュメモリ部14からデータを読み出し、オリジナルのデータセットが格納されている磁気ディスク装置5に繋がるディスクIF部12へ転送し、そこから磁気ディスク装置5に転送し、書き込ませる。

【0107】ディスクアレイ制御ユニット1-1-1のチャネルIF部11内のマイクロプロセッサ101は、オリジナルのデータセットのデータ更新を監視しており、ディスクアレイ制御ユニット1-1-1の共有メモリ部13内にオリジナルのデータセットのデータ更新量を示す制御情報を格納する。データ更新量があらかじめ定められた値以上になると、前記のマイクロプロセッサ101は、オリジナルのデータセットの更新内容を複製のデータセットに反映するように、オリジナルのデータセットが格納されている磁気ディスク装置5が繋がるデ

(単3) 01-256003 (P2001-0003)

ィスクIF部12内のマイクロプロセッサ101に命令を発行する。それを受け、マイクロプロセッサ101は、磁気ディスク装置5から更新されたデータを読み出し、その更新データのアドレスを複製のデータセットのアドレスに変換する。その更新データを、ディスクアレイ制御ユニット1-1-1のCM-SW111、筐体間CM-SW122、及びディスクアレイ制御ユニット1-1-2のCM-SW111を介して、ディスクアレイ制御ユニット1-1-2のキャッシュメモリ部14へ転送し、書き込む。次に、そのキャッシュメモリ部14から更新データを読み出し、複製のデータセットが格納されている磁気ディスク装置5に繋がるディスクIF部12へ転送し、そこから磁気ディスク装置5に転送し、書きこむ。

【0108】上記の動作により、通常業務において、オリジナルのデータセットと複製のデータセットを保持する。

【0109】ディスクアレイ制御ユニット1-1-2に繋がるホストコンピュータ50から、目的のデータセットに対してバックアップをとる要求があった場合、そのホストコンピュータ50が繋がるチャネルIF部11内のマイクロプロセッサ101は、ディスクアレイ制御ユニット1-1-1の共有メモリ部13を介して、通常業務を行っているホストコンピュータ50が繋がるチャネルIF部11内のマイクロプロセッサ101に複製のデータセットに対するデータの更新を一時中断するように命令を発行する。それを受けたマイクロプロセッサ101は、データの更新を一時中断する。次に、バックアップの要求を出したホストコンピュータ50に繋がるチャネルIF部11内のマイクロプロセッサ101は、複製のデータセットが格納されている磁気ディスク装置5に繋がるディスクIF部12内のマイクロプロセッサ101に、ディスクアレイ制御ユニット1-1-2の共有メモリ部13を介して命令を発行し、複製のデータセットを磁気ディスク装置5から読み出し、ディスクIF部12へ転送し、ディスクIF部12からディスクアレイ制御ユニット1-1-2のキャッシュメモリ部14に転送し、書き込ませる。それが終了すると、前記のチャネルIF部11内のマイクロプロセッサ101は、ディスクアレイ制御ユニット1-1-2のキャッシュメモリ部14から複製のデータセットを読み出し、チャネルIF部11へ転送し、そこからバックアップの要求を出したホストコンピュータ50へ送る。

【0110】データセットのバックアップが終了すると、バックアップの要求を出したホストコンピュータ50に繋がるチャネルIF部11内のマイクロプロセッサ101は、通常業務を行うホストコンピュータ50が繋がるチャネルIF部11内のマイクロプロセッサ101に、ディスクアレイ制御ユニット1-1-1の共有メモリ部13を介して命令を発行し、オリジナルのデータセ

ット内のバックアップ処理中に更新されたデータを、複製のデータセットに反映させる。この方法は、上記の通常業務における更新データの反映を行う方法と同様である。本例によれば、上記機能を実現する場合、内部の相互結合網及びキャッシュメモリ部14を介して、2つのディスクアレイ制御ユニット1-1-1、1-1-2間でデータの移行を行えるため、両方のディスクアレイ制御ユニット1-1-1、1-1-2のチャネルIF部を介してデータを移行する必要が無くなり、上記機能を実行しているときの性能低下を抑えることが可能となる。したがって、ユーザの通常業務の効率を低下させることができなくなる。

【0111】実施例2及び実施例3の構成のディスクアレイ制御装置1においても、本実施例を実施する上で問題はなく、本実施例と同様の効果が得られる。

【0112】その他の利用例として以下のものがある。実施例1、実施例2、及び実施例3に示すディスクアレイ制御装置1では、ホストコンピュータ50からディスクアレイ制御ユニットに、例えばデータのリード要求があり、そのディスクアレイ制御ユニットに接続された磁気ディスク装置5にデータが無かった場合、相互結合網を介して、該当するデータを格納している磁気ディスク装置5が接続されたディスクアレイ制御ユニットからデータを読み出し、ホストコンピュータ50へ要求データを返す必要がある。このようにディスクアレイ制御ユニット間を跨ってデータを移行させることによりデータのリード/ライトを行う場合は、そうでない場合に比べて性能が低下する。

【0113】上記のようなデータ移行を抑えるため、あるホストコンピュータ50からアクセスされる頻度の高いデータの中で、該ホストコンピュータ50が繋がるディスクアレイ制御ユニットとは異なるディスクアレイ制御ユニットに接続された磁気ディスク装置5に格納されているデータを、該ホストコンピュータ50が繋がるディスクアレイ制御ユニットに接続された磁気ディスク装置5に移行する機能が必要となる。

【0114】実施例1に示すディスクアレイ制御装置1において、上記のデータ移行の方法を図8を用いて説明する。

【0115】チャネルIF部11内のマイクロプロセッサ101は、全ての磁気ディスク装置5内のデータセット(論理ボリュームに対応)へのアクセス頻度を監視しており、自身と同じディスクアレイ制御ユニット1-1-1内の共有メモリ部13内に前記のデータセットのアクセス頻度を示す制御情報を格納する。

【0116】ここで、ディスクアレイ制御ユニット1-1-1に繋がるホストコンピュータ50から、ディスクアレイ制御ユニット1-1-2に繋がる磁気ディスク装置5内のデータセットにアクセスが集中し、アクセス頻度があらかじめ定められた値を超えた場合、ディスクア

(単4) 01-256003 (P2001-(\*03)

レイ制御ユニット1-1-1のチャネルIF部11内のマイクロプロセッサ101は、該当するデータセットを格納している磁気ディスク装置5が繋がるディスクIF部12内のマイクロプロセッサ101に、ディスクアレイ制御ユニット1-1-2内の共有メモリ部13を介して命令を発行し、該当するデータセットを読み出し、ディスクアレイ制御ユニット1-1-2のキャッシュメモリ部14へ転送し、書き込ませる。

【0117】次に、ディスクアレイ制御ユニット1-1-1のチャネルIF部11内のマイクロプロセッサ101は、ディスクアレイ制御ユニット1-1-2のキャッシュメモリ部14から該当するデータを読み出し、ディスクアレイ制御ユニット1-1-1のキャッシュメモリ部14に転送する。次に、ディスクアレイ制御ユニット1-1-1のディスクIF部12内のマイクロプロセッサ101に、ディスクアレイ制御ユニット1-1-1内の共有メモリ部13を介して命令を発行し、ディスクアレイ制御ユニット1-1-1のキャッシュメモリ部14から該当するデータを読み出し、磁気ディスク装置5に書き込ませる。本例によれば、上記のようなデータ移行を行う場合、内部の相互結合網及びキャッシュメモリ部14を介して、2つのディスクアレイ制御ユニット1-1間でデータ移行を行えるため、両方のディスクアレイ制御装置のチャネルIF部を介してデータを移行する必要が無くなり、上記データ移行を実行しているときの性能低下を抑えることが可能となる。したがって、ユーザの通常業務の効率を低下させることがなくなる。

【0118】実施例2及び実施例3の構成のディスクアレイ制御装置においても、本実施例を実施する上で問題はなく、本実施例と同様の効果が得られる。

【0119】次に実施例1、2、3の実装例について述べる。図14は、実施例1の図7に示すディスクアレイ制御ユニット1-1を筐体201に搭載した一例を示している。

【0120】図7に示すチャネルIF部11はチャネルIFパッケージ(PK)311上に実装され、ディスクIF部12はディスクIFPK312上に実装され、SM-SW110およびCM-SW111はスイッチPK320上に実装され、共有メモリ部13及びキャッシュメモリ部14はメモリPK330上に実装されている。また、バックプレーン340上にはSMアクセスバス135、136、及びCMアクセスバス137、138が配線されており、バックプレーン340に上記各PKを挿す形態となっている。

【0121】スイッチPK320には、筐体間SMバス141用のケーブルと筐体間CMバス142用のケーブルが接続され、それぞれのケーブルの他端は、筐体201の側面にあるコネクタ221、222にそれぞれ接続される。ケーブルの図示は省略してある。350は電源ボックスであり、上記のPKに電力を供給する。このよ

うにディスクアレイ制御ユニットは1つの筐体としてディスクアレイ制御装置の機能を備えている。

【0122】図15は、図14に示す筐体201を2台接続する場合の一例を示している。

【0123】スイッチボックス210には、図8に示す筐体間SM-SW121、及び筐体間CM-SW122が搭載されている。筐体間SM-SW121に繋がる筐体間SMバス141はコネクタ221に、筐体間CM-SW122に繋がる筐体間CMバス142はコネクタ222に接続されている。

【0124】2台の筐体201、202を接続する場合は、筐体201の筐体間SMバス用コネクタ221とスイッチボックス210のコネクタ221をケーブル231で、筐体201の筐体間CMバス用コネクタ222とスイッチボックス210のコネクタ222をケーブル232で繋ぐ。同様に、筐体202のコネクタ221とスイッチボックス210のコネクタ221をケーブル231で、筐体202のコネクタ222とスイッチボックス210のコネクタ222をケーブル232で繋ぐ。

【0125】上記のようにすることで、1つの筐体でサポートできないホストコンピュータへの接続チャネル数、あるいは1つの筐体でサポートできない記憶容量をサポートすることが可能となる。

【0126】ここで、筐体201を基本の筐体、筐体202を拡張用の筐体とし、スイッチボックス210を筐体202の中に搭載することも出来る。こうすることにより、基本の筐体201の製造コストを上げることなく、スイッチボックス210の設置スペースを削除することができる。

【0127】実施例2、または実施例3のディスクアレイ制御装置に本実施例を適用することは、何の問題もない。

【0128】図18に筐体間SM-SW121および筐体間CM-SW122で3台のディスクアレイ制御ユニットを接続する場合の接続形態を示す。このとき各SWは前述したように接続台数が増えているので2台のディスクアレイ制御ユニットを接続する場合より容量は大きいものが必要とされる。図示のごとく各ディスクアレイ制御ユニット1-1-1～1-1-3は筐体間SMバス141および筐体間CMバス142によって、筐体間SM-SW121および筐体間CM-SW122にそれぞれ接続され、全体として1つのディスクアレイ制御装置として機能する。

【0129】図19にその実装例を示す。ここではスイッチボックス210は別筐体として実装されている。これに筐体間SMバス用ケーブル231と筐体間CMバス用ケーブル232を介してそれぞれコネクタ221、コネクタ222により筐体201、202、203がそれぞれ接続される。スイッチボックス210にディスクアレイ制御ユニットを4台以上接続する容量とコネクタを

(15) 101-256003 (P 2001-0" #03

用意しておけば後からの増設は容易である。

【0130】図21はスイッチボックス210を通るデータのフォーマットを示す。データはパケットの形態を取り、宛先アドレス401、コマンド部402、データ部403からなる。アドレスは共有メモリ、キャッシュメモリ上のアドレスである。

【0131】図22はスイッチボックス210内に設けられたスイッチングのためのスイッチ切り替えテーブル410を示す。ここには宛先アドレスとそのアドレスを含むディスクアレイ制御ユニットの番号の対応が記憶されている。スイッチボックス210はパケット400のアドレス401からこのスイッチ切り替えテーブルを参照し、スイッチの切り替え先を求め、スイッチの切り替え制御を行なう。

【0132】ディスクアレイ制御ユニットを増設する場合は次の手順による。スイッチボックス210にディスクアレイ制御ユニットを増設するコネクタに余分があれば、そのコネクタにケーブル231、ケーブル232を接続する。余分がなければスイッチボックスを多段に接続した上でそのコネクタにケーブル231、ケーブル232を接続する。それと共に、スイッチボックス210内のスイッチ切り替えテーブル410のアドレスとポートNo.を増設したディスクアレイ制御ユニット分だけ書き加える。前記のアドレスを予め書き込んでおき、増設した場合有効のフラグを立てるやり方もある。

【0133】図20は他の接続例を示すものである。図示のように3つのディスクアレイ制御ユニットが直列に接続されている。このときSM-SW110、CM-SW111は入力された情報をそのまま他のディスクアレイ制御ユニットに転送するブリッジ機能を持っている。図ではSM-SW110とCM-SW111を有しているが、その代わりにバス構造としても良い。そして、複数のディスクアレイ制御ユニットをバス同士接続した共通バスで結合させても良い。

【0134】次に更に他の実装例を示す。図16に示すように、図14に示す筐体201に搭載したディスクアレイ制御ユニット1-1のパッケージ(PK)の枚数を減らし、最小構成のディスクアレイ制御ユニット1-1を搭載した筐体205とする。

【0135】図17に示すように、1つの筐体206内に2個以上の筐体205とスイッチボックス210を搭載し、筐体205間を実施例6で示した方法と同様の方法で、スイッチボックス210を介して接続することにより、中規模から大規模な構成のディスクアレイ制御装置を構成することが可能となる。

【0136】なお、筐体205は今まで述べてきたディスクアレイ制御ユニットの機能を持っているものであれば良く、筐体の形をとらずにモジュールと呼ばれる形態をとっても良い。また、図17の場合ディスクアレイ制御ユニット毎に電源ボックスを持たせるか、共通の電

源ボックスから給電するかは実装上の事柄として適宜決められるものである。

【0137】

【発明の効果】本発明によれば、複数台のディスクアレイ制御装置を1つのディスクアレイ制御装置として運用しようとする場合、複数のディスクアレイ制御装置間でのデータ移行による性能低下を抑えるディスクアレイシステムを提供することが可能となる。

【図面の簡単な説明】

【図1】本発明によるディスクアレイ制御装置の構成を示す図。

【図2】従来のディスクアレイ制御装置の構成を示す図。

【図3】従来のディスクアレイ制御装置の他の構成を示す図。

【図4】従来のディスクアレイ制御装置の他の構成を示す図。

【図5】本発明によるディスクアレイ制御装置の他の構成を示す図。

【図6】本発明によるディスクアレイ制御装置の他の構成を示す図。

【図7】図1に示すディスクアレイ制御ユニット内の詳細構成を示す図。

【図8】図7に示すディスクアレイ制御ユニットを複数台接続する構成を示す図。

【図9】図5に示すディスクアレイ制御ユニット内の詳細構成を示す図。

【図10】図9に示すディスクアレイ制御ユニットを複数台接続する構成を示す図。

【図11】図6に示すディスクアレイ制御ユニット内の詳細構成を示す図。

【図12】図11に示すディスクアレイ制御ユニットを複数台接続する構成を示す図。

【図13】図7に示すディスクアレイ制御ユニットを複数台接続する他の構成を示す図。

【図14】本発明によるディスクアレイ制御ユニットの筐体への搭載例を示す図。

【図15】本発明によるディスクアレイ制御ユニットを搭載した筐体を複数台接続する構成を示す図。

【図16】本発明によるディスクアレイ制御ユニットの筐体への搭載の他の例を示す図。

【図17】本発明によるディスクアレイ制御ユニットを複数台、1つの筐体へ搭載する例を示す図。

【図18】筐体間スイッチにより3台のディスクアレイ制御システムを接続する配線構造を示す図。

【図19】図18の配線構造の一実装例を示す図。

【図20】本発明によるディスクアレイ制御ユニットを3台以上接続するための他の実施例を示す図。

【図21】スイッチボックスに与えられる情報のデータフォーマットの一例を示す図。

(単6) 01-256003 (P2001-0複雑)

【図22】スイッチボックス内に設けられたスイッチ切り替えテーブルの一例を示す図。

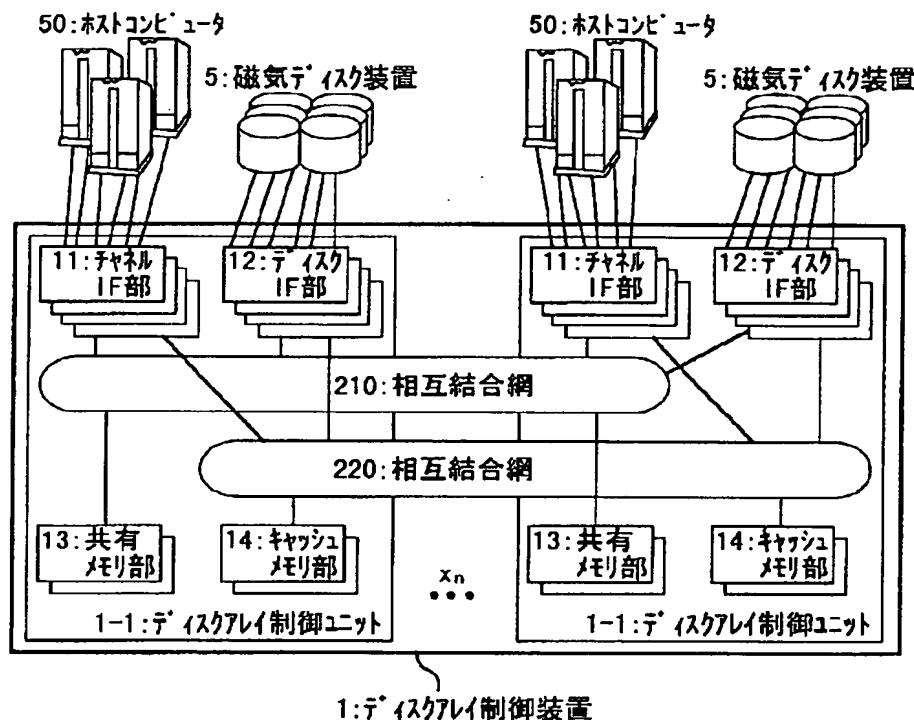
【符号の説明】

1:ディスクアレイ制御装置、1-1…ディスクアレイ

制御ユニット、5…磁気ディスク装置、11…チャネルIF部、12…ディスクIF部、13…共有メモリ部、14…キャッシュメモリ部、210、220…相互結合網、50…ホストコンピュータ

【図1】

図 1

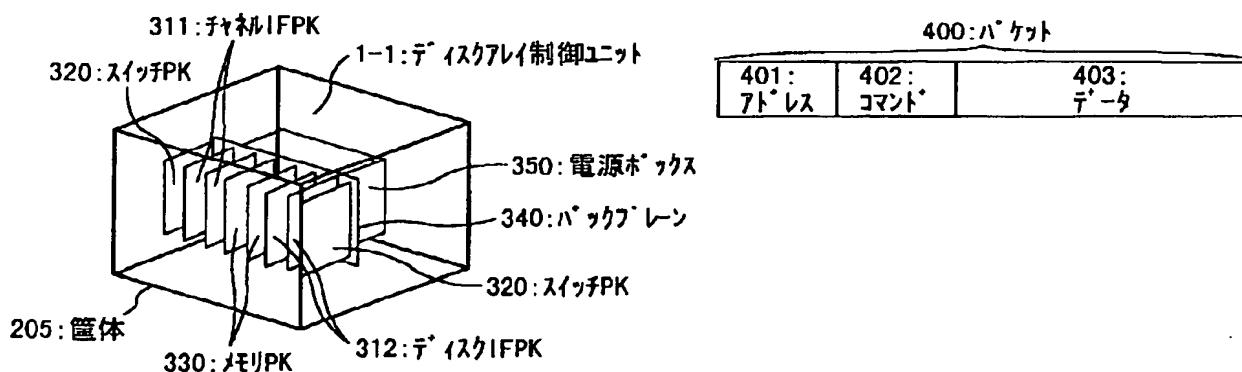


【図16】

図 16

【図21】

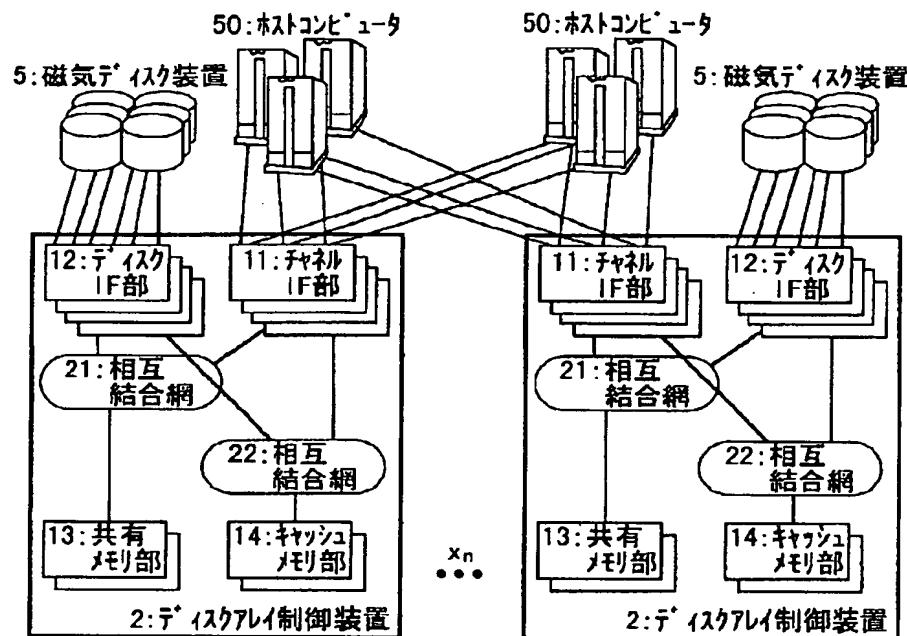
図 21



(17) 01-256003 (P2001-F03)

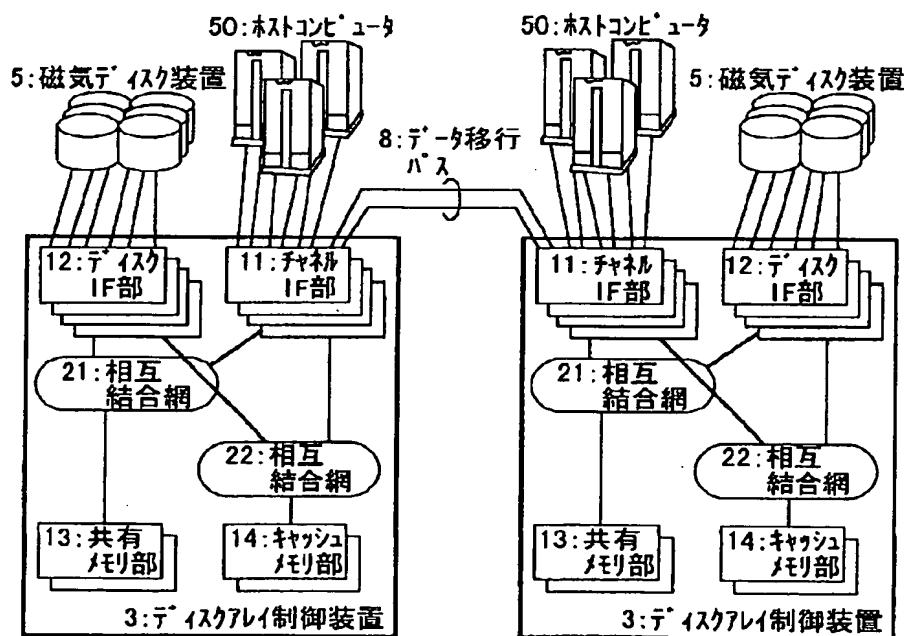
【図2】

図 2



【図3】

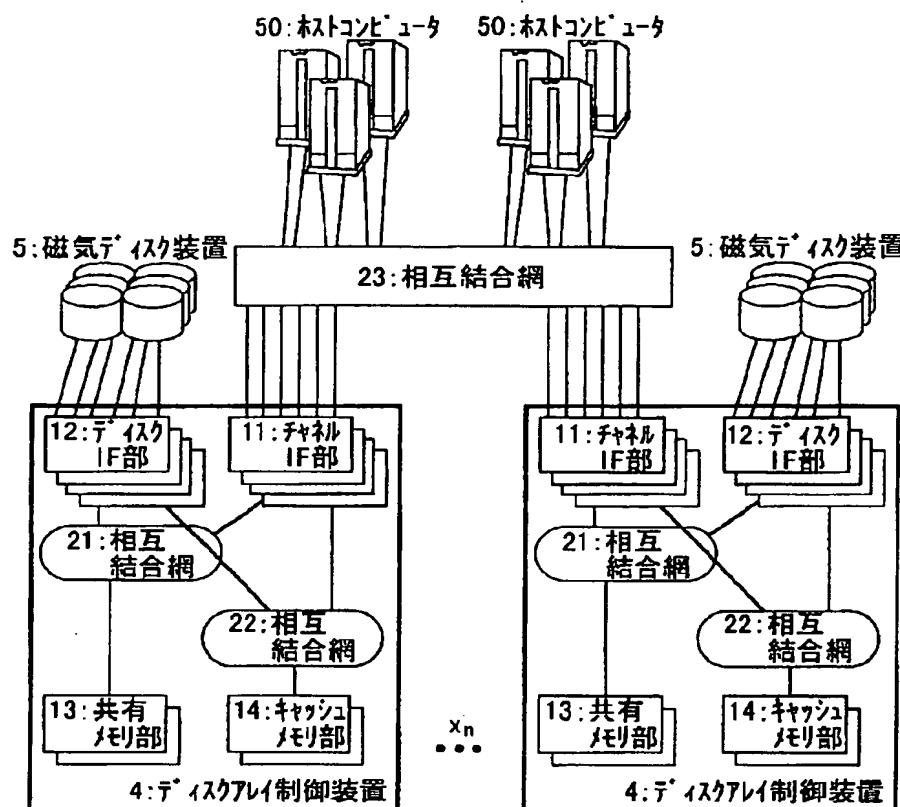
図 3



(単8) 01-256003 (P2001-u雑誌)

【図4】

図 4



【図22】

図 22

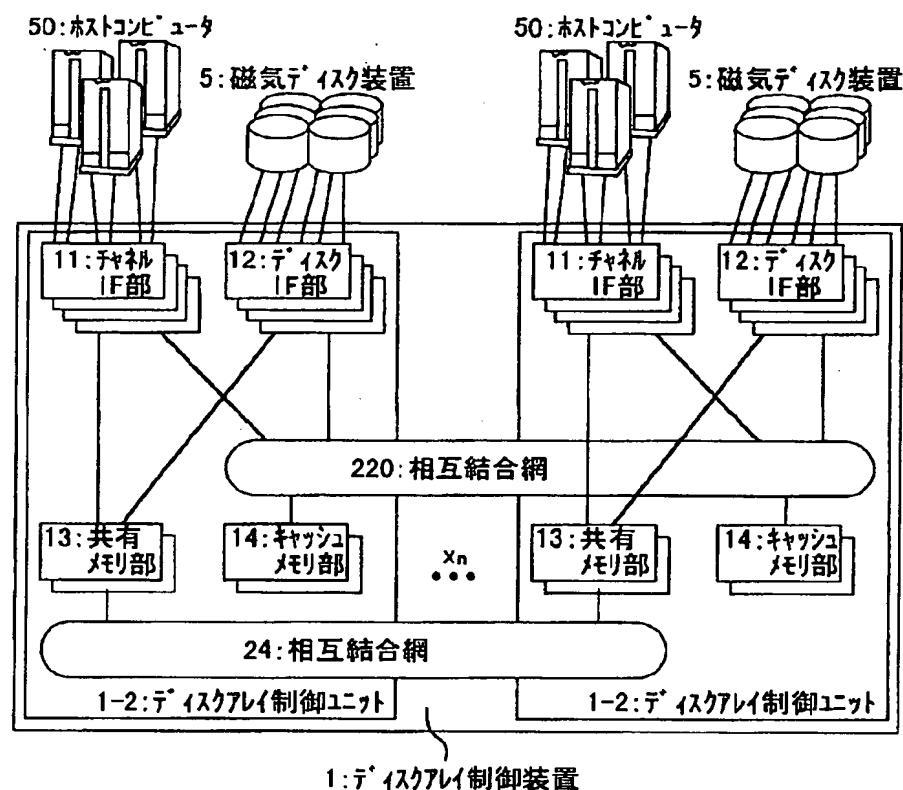
## 410:スイッチ切替テーブル

アドレス	ポートNo
0x0...000	0
0x0...001	0
⋮	⋮
0x0...OFF	0
0x0...100	1
⋮	⋮
0x0...200	2
⋮	⋮

(19) 01-256003 (P2001-7'03)

【図5】

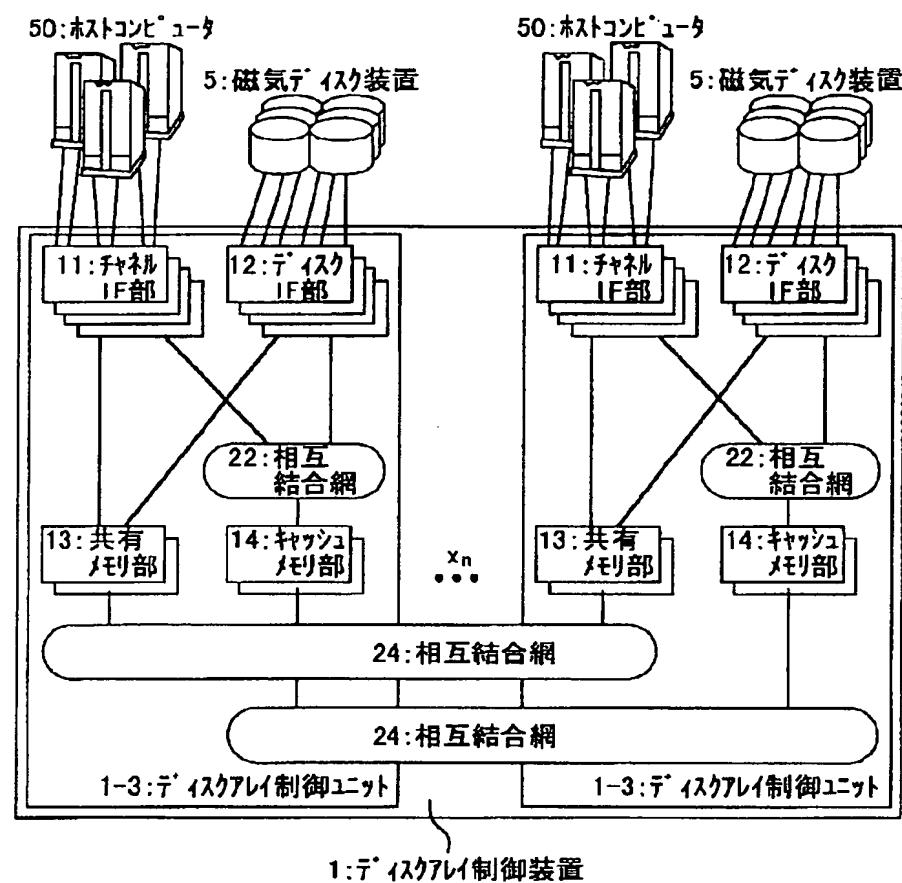
図 5



(20) 01-256003 (P2001-, 姉

【図6】

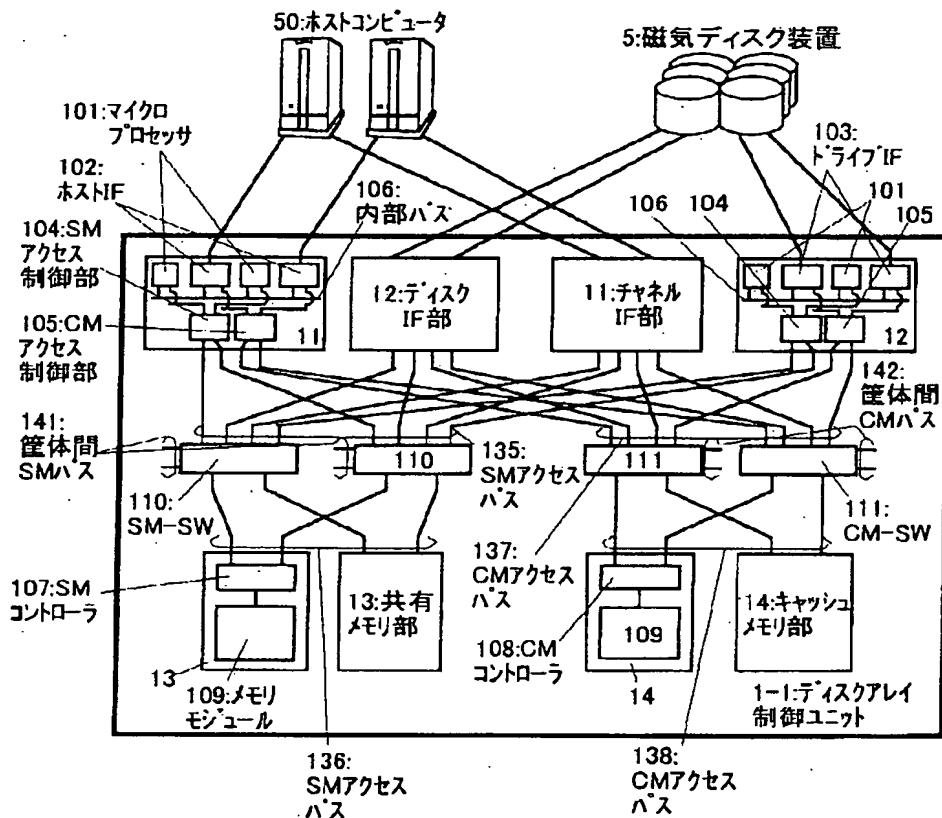
図 6



(21) 01-256003 (P 2001-103)

【図7】

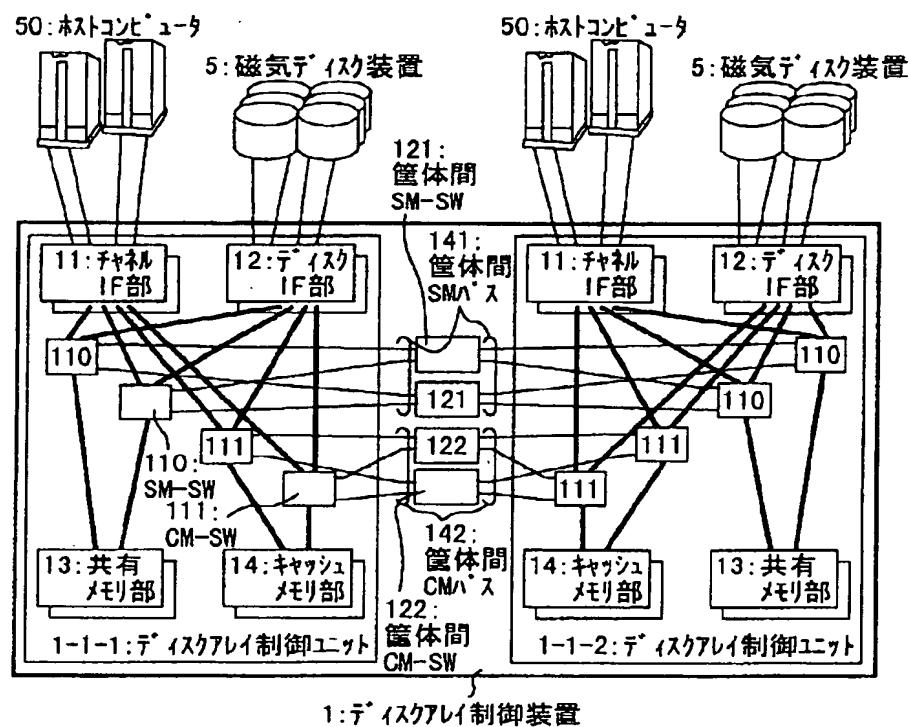
図7



(22) 01-256003 (P2001-ch; 菊娃

【図8】

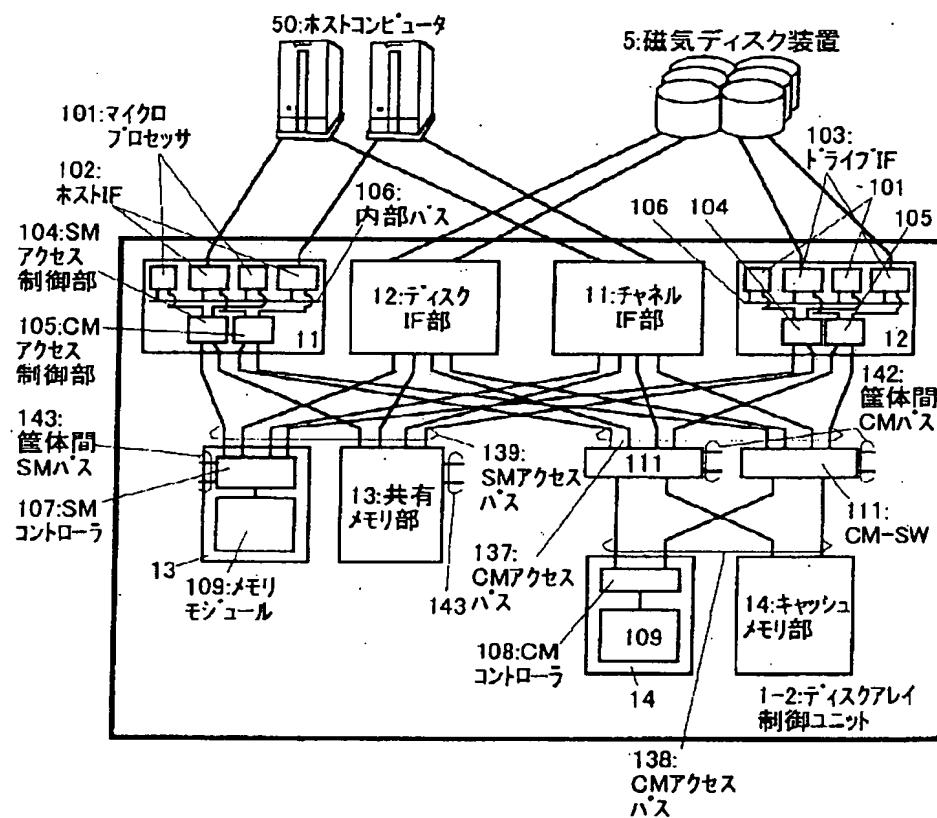
図 8



(23) 101-256003 (P2001-9L03)

【図9】

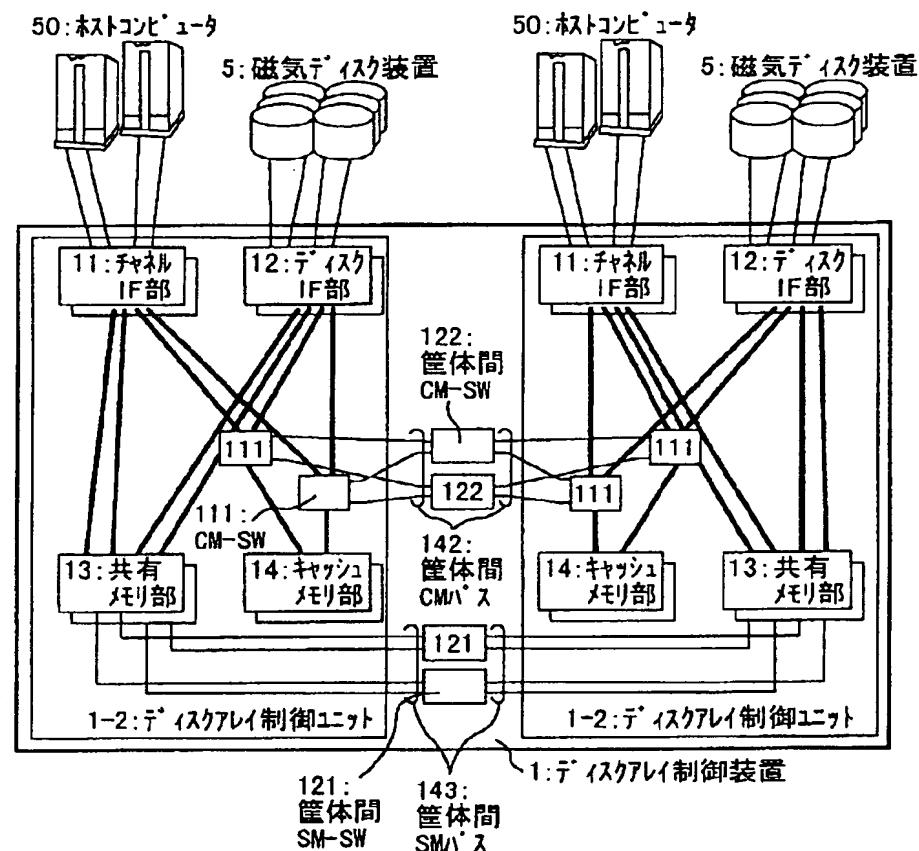
図9



(24) 01-256003 (P2001-903)

【図10】

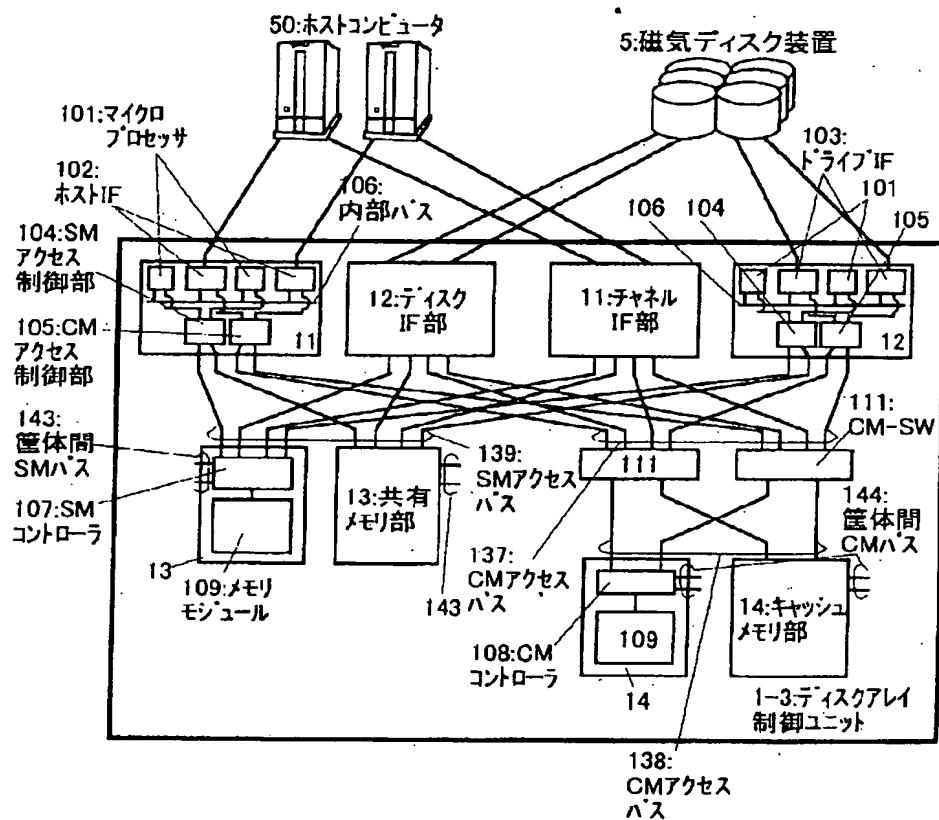
図 10



(25) 101-256003 (P2001-8雛姫)

【図11】

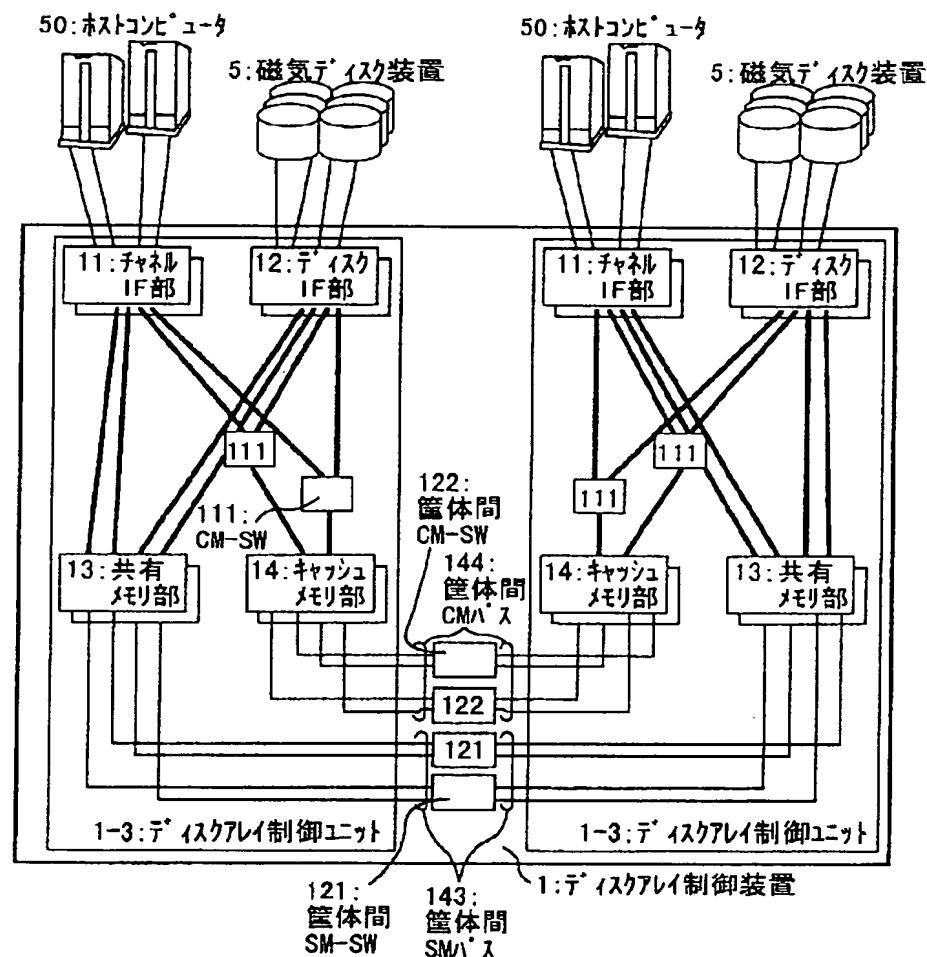
図11



(26) 01-256003 (P2001-E903)

【図12】

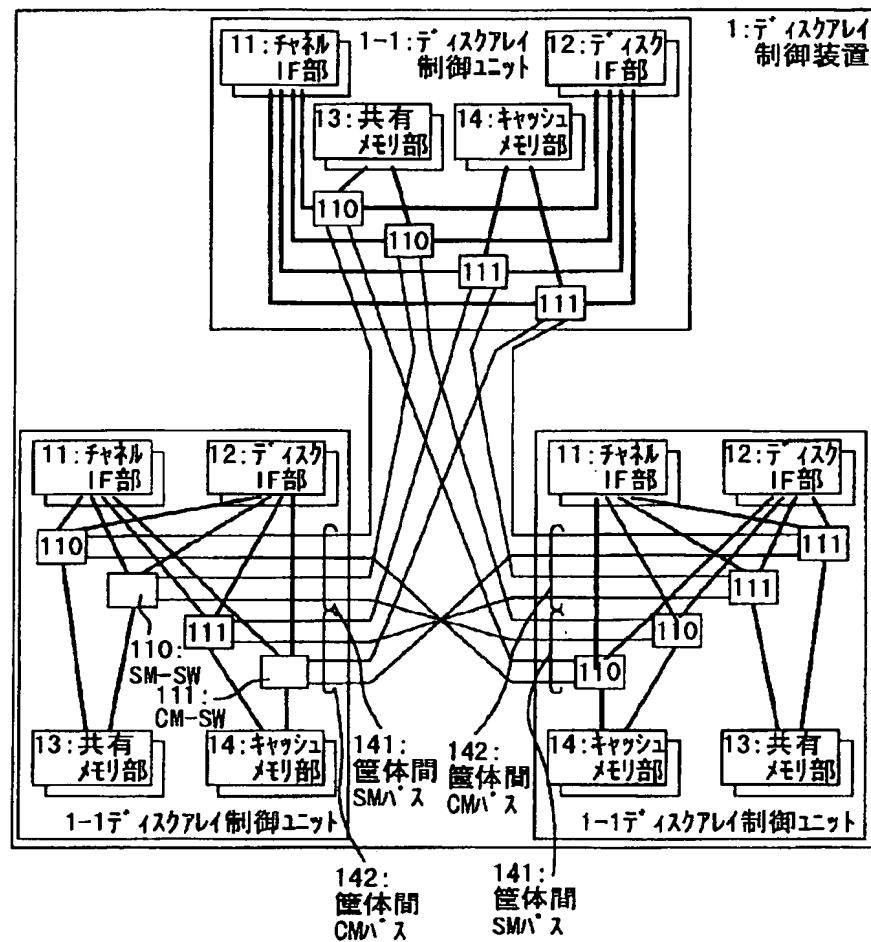
図 12



(27) 01-256003 (P2001-E:=03

【図13】

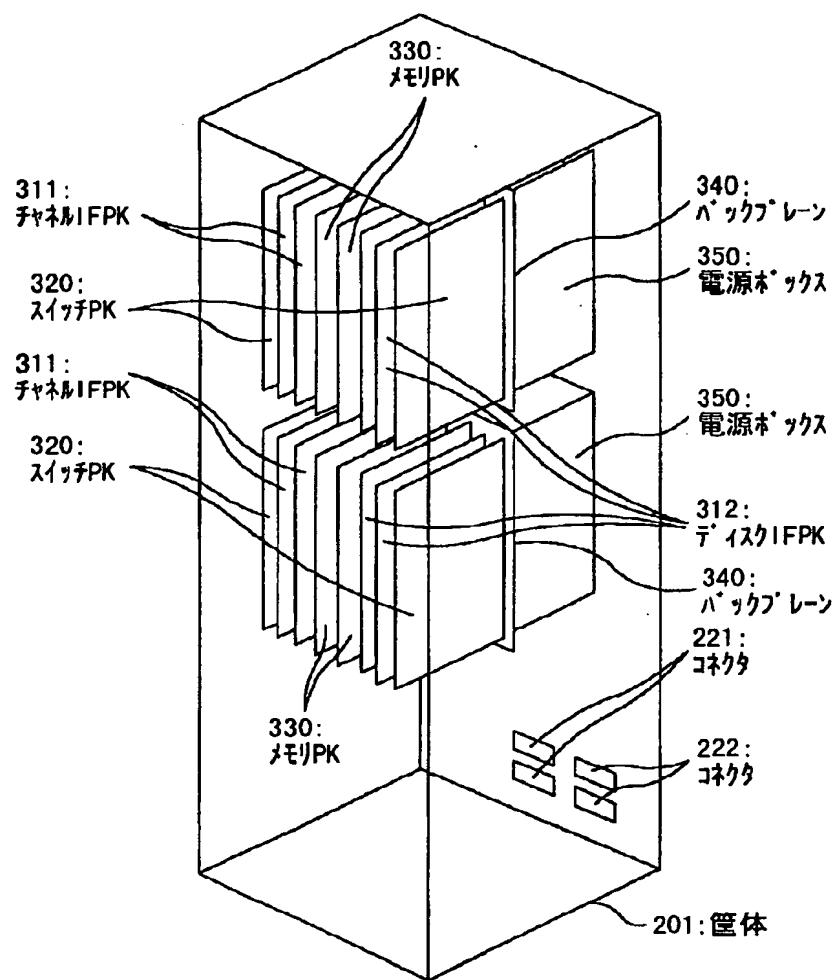
図 13



(28) 01-256003 (P2001-u9械娃

【図14】

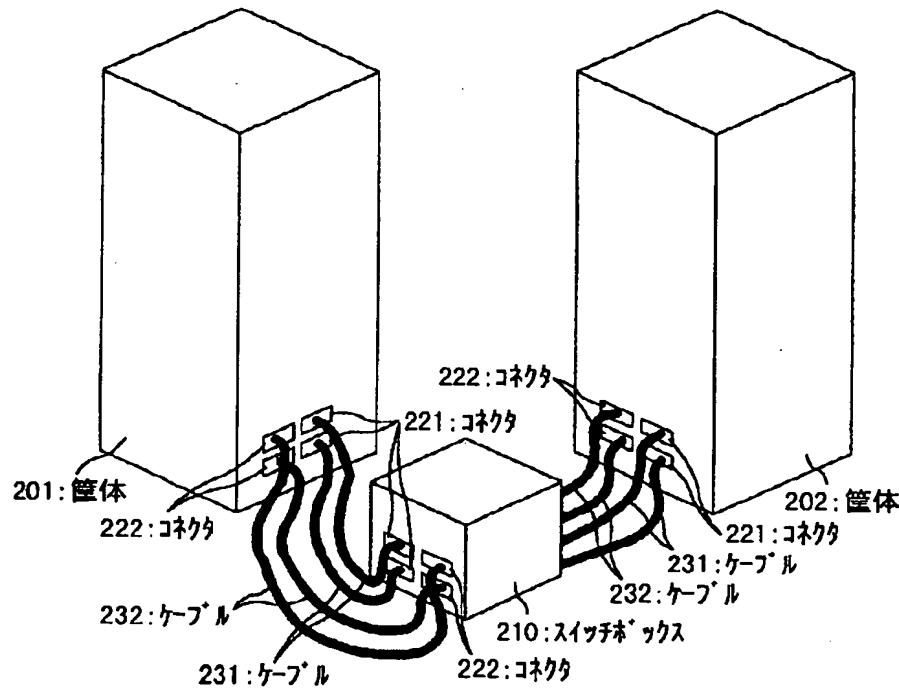
図 14



(29) 01-256003 (P2001-003

【図15】

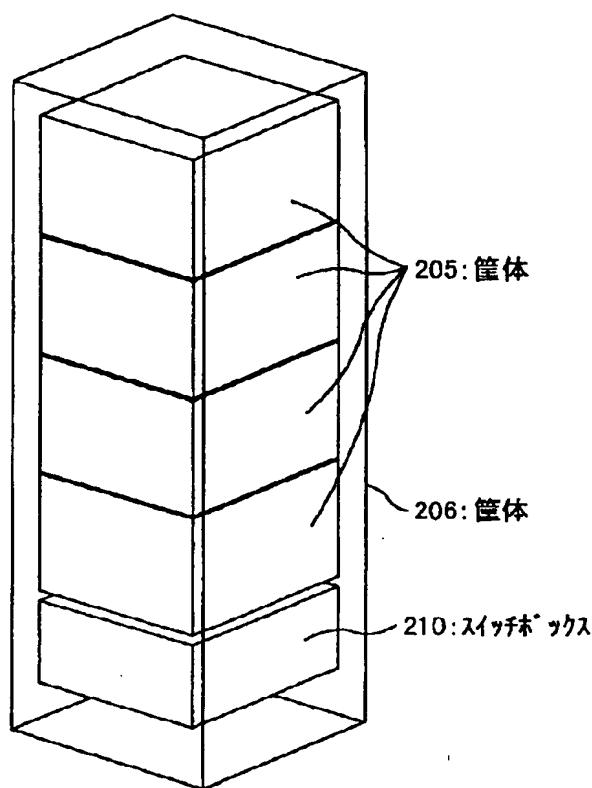
図 15



(B0) 01-256003 (P2001-}03

【図17】

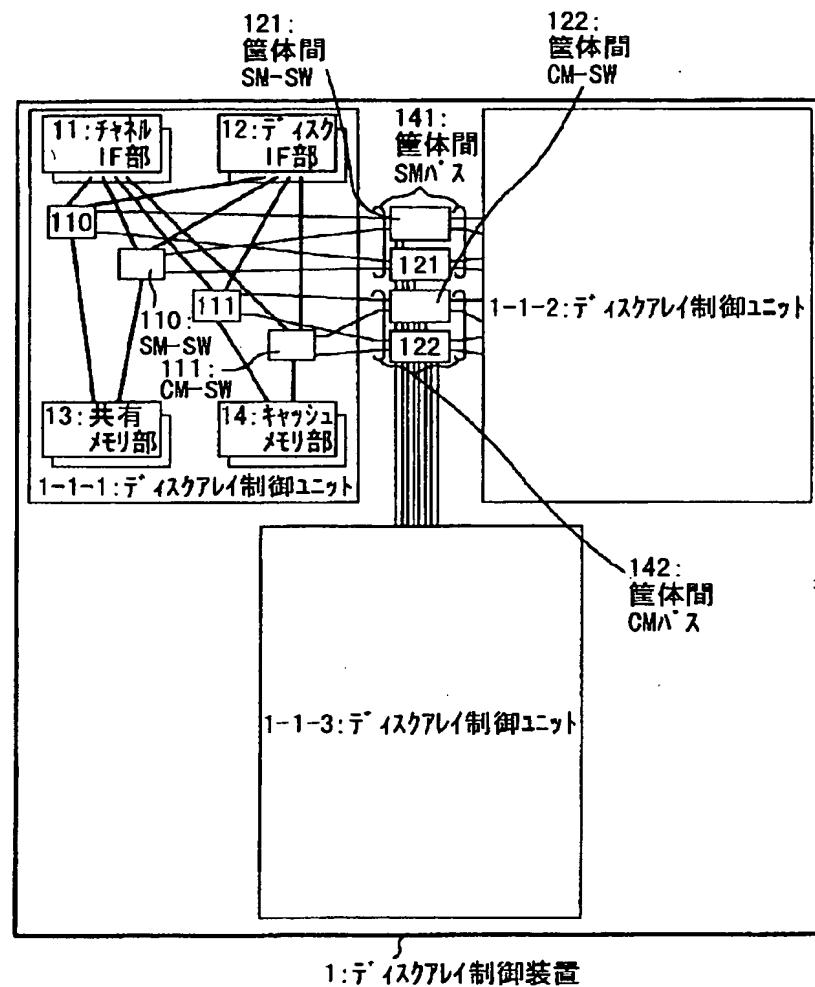
図 17



(S1) 01-256003 (P2001-娃

【図18】

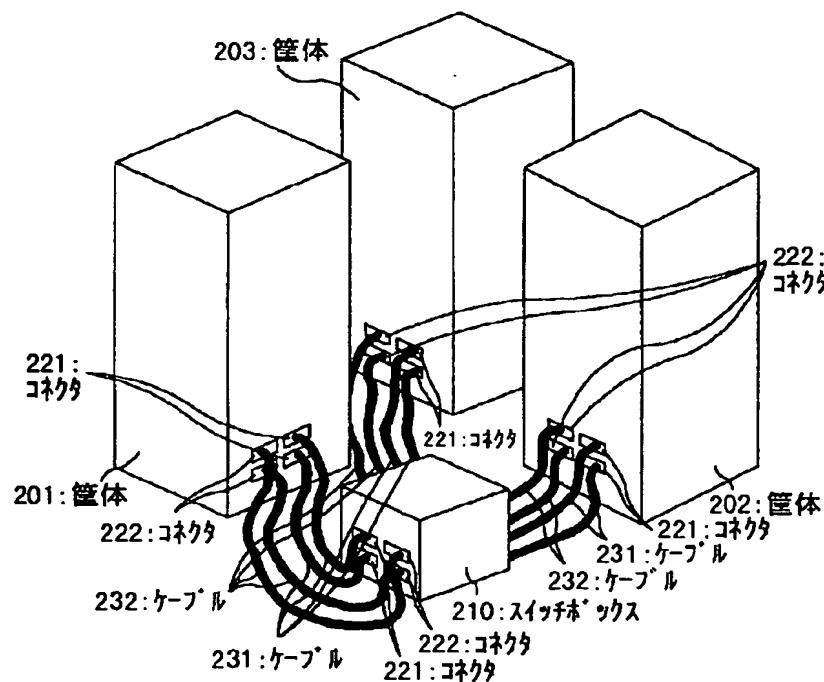
図 18



(B2) 01-256003 (P2001-%遙蛙

【図19】

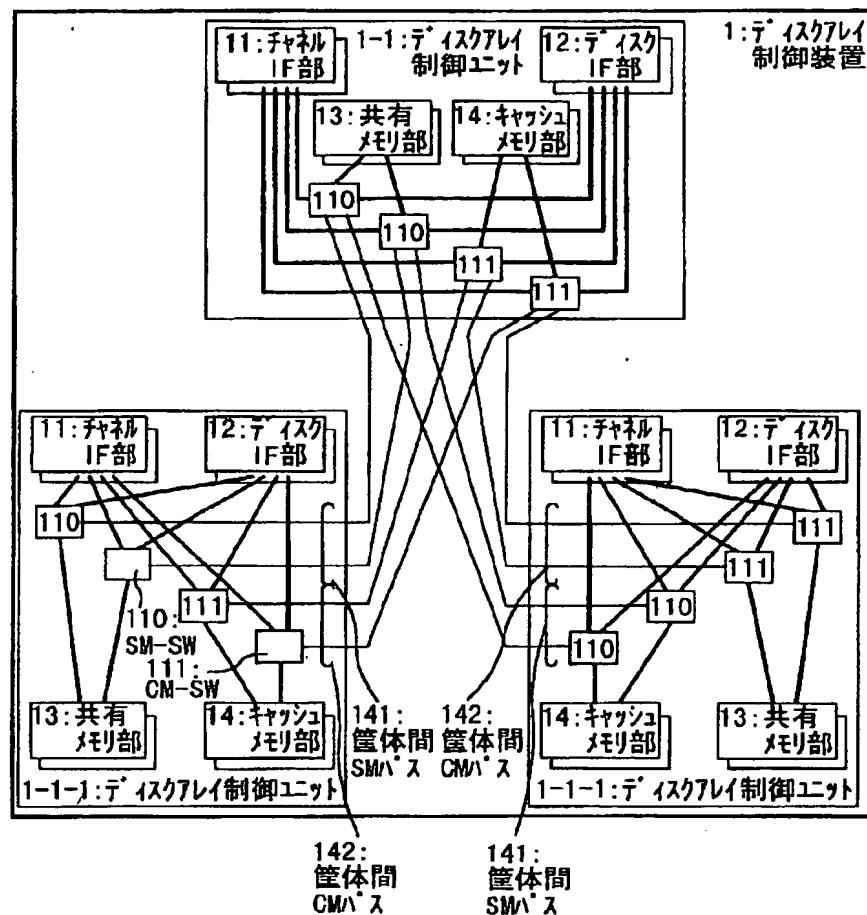
図 19



(33) 01-256003 (P 2001-")娃

【図20】

図 20



フロントページの続き

(51) Int.Cl. 7  
G 06 F 12/16識別記号  
320F I  
G 06 F 12/16

テマコード(参考)

320 L

(72) 発明者 藤林 昭  
東京都国分寺市東恋ヶ窪一丁目280番地  
株式会社日立製作所中央研究所内(72) 発明者 櫻井 直  
神奈川県小田原市国府津2880番地 株式会  
社日立製作所ストレージシステム事業部内  
F ターム(参考) 5B005 JJ01 MM12 WW02 WW04  
5B018 GA04 HA04 MA14  
5B065 BA01 CA30 CE22 CE26 CH01  
CH13